ARO 25839.12-MA

AD-A212 097

SMALL SAMPLE THEORY
FOR STEADY STATE CONFIDENCE INTERVALS

by

Chia-Hon Chien

TECHNICAL REPORT No. 37

June 1989

DTIC
ELECTE
SEP 12 1989
S B D

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

Contents

List of Tables

Abstract

The goal of this dissertation is to develop a nonparametric method for obtaining a confidence interval for the mean of a stationary sequence. As indicated in the literature, nonparametric confidence intervals in practice often have undesirable small-sample asymmetry and coverage characteristics. These phenomena are partially due to the fact that the third and fourth cumulants of the point estimator for the stationary mean, unlike those of the standard normal random variable, are not zero. We will apply Edgeworth and Cornish-Fisher expansions to obtain asymptotic expansions for the errors associated with confidence intervals. The analysis isolates various elements that contribute to errors and makes it possible for us to estimate each element and hopefully correct the errors to a smaller order. We will use Glynn's method to develop first and second order pivots for the confidence intervals. Furthermore, these procedures also improve the asymptotic order of confidence interval accuracy.

Keywords: simulation, steady-state, nonparametric confidence intervals, Edgeworth expansions, stationary processes

Chapter 1

Introduction

The use of stochastic analysis has long been an important tool in the study of complex systems such as manufacturing, computer, and communications systems. As the complexity of systems grow, it appears to be the rule rather than the exception that many detailed stochastic models are so complex such that it is extremely difficult or impossible to obtain an exact analytic solution. Although simulation was often viewed as a "method of last resort" to be used only when everything else failed, the intrinsic complexity of stochastic models for real world systems, recent advances in simulation methodology, and the availability of computing power and software packages have made simulation one of the most widely used tools in system analysis and operations research.

In a typical simulation study, we first construct a mathematical model that corresponds to the real world system to be studied. We then perform computer sampling experiments on the model, collect and analyze the output sequences, and make inferences about the behavior of the system. Since simulation is a statistics based computer sampling procedure, appropriate statistical techniques must be used in both design and analysis of the simulation study before any meaningful conclusion can be drawn.

Simulation output analysis is an area of active research. In most situations, we want to estimate quantities associated with the stochastic system being simulated. When estimating a quantity only by a sample mean, no indication of the

variability of the estimate is given. The standard deviation of the estimate is one device used for indicating the reliability or precision of an estimate. What is more informative, however, is a confidence interval. Basically, we first obtain a point estimate, calculate its standard deviation, and then construct a confidence intervals for each quantity to access the statistical variability of the point estimate. It is generally thought that the construction of confidence intervals is one of the most difficult and important issues in the field of simulation output analysis.

Although there is a large statistical literature devoted to the assignment of confidence intervals, most of it can not be directly applied to simulation output analysis. One important reason for this is that most stochastic processes associated with real world simulation studies do not satisfy the standard assumptions in the statistical literature, namely, independence, stationarity, etc. At present, there are many methods proposed in the simulation literature. The most common methods are replications, batch means, overlapping batch means, regenerative, spectral, autoregressive, autoregressive moving average, and standardized time series.

On the other hand, the recognition that computing costs are decreasing has allowed statisticians and simulation researchers to consider confidence intervals methods which are computationally more intensive but statistically better behaved than previous techniques. Among these methods are the jackknife, bootstrap, other resampling plans, and Johnson-Glynn pivital tranformations.

The underlying mathematical model in this study is a stationary stochastic process which satisfies certain regularity conditions. As in any steady state analysis of time series, a key assumption is that the initial state should have a very small effect on the overall behavior of the system. However, when studying a specific real world system, this assumption does not always hold. On the other hand,

initial effects decay exponentially over time for several classes of stochastic processes so that steady-state simulation is often appropriate for long simulation runs. In Chapter 5, we give several numerical examples showing that their initial conditions do not affect the behavior of the confidence intervals.

In this dissertation we develop a nonparametric method for obtaining a more accurate asymptotic confidence interval for the sample mean of a stationary process which satisfies certain regularity conditions. By "more accurate asymptotic confidence interval" we mean that the coverage error, which is the difference between the nominal and the actual coverage, associated with our confidence interval will be of lower order than that of the traditional methods.

Our starting point is the traditional batch means method. We use the idea of Johnson-Glynn pivotal transformations to obtain better confidence intervals for the quantities of interest. No assumption has been made that the observed data are sampled from either i.i.d., regenerative, or ARMA processes. We believe this more general case is robust for the output analysis of real world simulation experiments. The procedures we propose do not require the selection of any critical constants that can not be reasonably preset. This fact and the less restrictive nature of the assumptions will also allow us to implement these procedures as a software package for the output analysis of many real world simulation studies.

The basic approach of the procedures we propose follow from Johnson [15], Glynn [12], and Titus [25]. As indicated in Glynn [12], nonparametric confidence intervals in practice often have undesirable small-sample asymmetry and coverage characteristics. We will apply Edgeworth expansion theory to obtain asymptotic expansions for the errors associated with confidence intervals. The analysis isolates the various elements that contribute to the errors. We then use Glynn's method to

develop first and second order pivots to the confidence intervals, which deal with asymmetry problems and coverage difficulties, respectively. These procedures also improve the asymptotic order of confidence interval accuracy in the sense that the actual coverage of the corrected confidence interval is closer to the nominal coverage rate than previous methods.

Johnson [15] is perhaps the first author to use these procedures. He derives a first order pivot for the t-statistic for independent and identically distributed samples. Glynn [12] extends this idea to a second order pivot for the ratio estimators of the regenerative processes. Titus [25] applies the same idea to asymptotically stationary autoregressive processes of finite order.

The organization of this dissertation is as follows. In Chapter 2, we develop the necessary background for cumulants, Edgeworth expansions, and Cornish-Fisher expansions. Some properties of the cumulants for a stationary process satisfying some regularity conditions are also discussed there. In Chapter 3, we derive the first and second order Johnson-Glynn pivotal transformations to correct the error in confidence interval coverage. Some computational and theoretical issues are also discussed there. To demonstrate how our method works, several numerical examples are displayed in Chapter 4. Chapter 5 summarizes the strength and weakness of the Johnson-Glynn pivots as applied to batch means method. The Appendix contains some of the more technical proofs.

Chapter 2

Mathematical Background

In this chapter, we will review the necessary mathematical background required for this dissertation. In Section 2.1, definition of cumulants as well as some important properties of cumulants will be given. Section 2.2 concerns the Edgeworth expansion while the Cornish-Fisher expansion is discussed in Section 2.3. In Section 2.4, some new uniqueness properties of the Cornish-Fisher expansion will be presented. Section 2.5 studies some properties associated with the cumulants of stationary processes. Finally, in Section 2.6, we will present results about the consistency properties of the sample moments for stationary processes.

2.1. Cumulants

Consider a random variable X with distribution function F_X. The moments $\mu'_r = EX^r$, $r \geq 0$, and central moments $\mu_r = E(X - EX)^r$, $r \geq 0$, are useful constants for measuring properties of X and, in some cases, uniquely characterize the distribution function F_X (Chung [8], pp. 98–99). For some statistical analyses another sequence of constants, the cumulants, are more useful from a theoretical point-of-view.

Let Φ_X be the characteristic function of X. Then, subject to existence, the cumulants of X, κ_r's, are formally defined by (Kendall and Stuart [16], p. 69)

$$\Phi_X(t) = \int_{-\infty}^{\infty} e^{itx} dF_X$$

5

$$= \sum_{r=0}^{\infty} \mu_r' \frac{(it)^r}{r!}$$

$$= \exp\left\{ \sum_{r=1}^{\infty} \kappa_r \frac{(it)^r}{r!} \right\} \tag{2.1}$$

Thus κ_r is the coefficient of $(it)^r/r!$ in $\log \Phi_X(\cdot)$ when the expansion in power series exists; $\log \Phi_X$ is called the *cumulant generating function* (c.g.f.) or the *second characteristic function*. However, the first terminology is somehow misleading in the sense that $\log \Phi_X$ exists in some neighborhood of the origin even if the moments and cumulants do not exist.

The proceeding expansions can be made rigorous. It is known that Φ_X is uniformly continuous in $(-\infty, \infty)$ and $\Phi_X(0) = 1$ (Chung [8], p. 143). Therefore there exists a neighborhood of the origin in which Φ_X is different from zero; let $|t| < \Delta$ be this neighborhood. By taking the principal branch of the logarithm, $\log \Phi_X$ can be uniquely defined for $|t| < \Delta$. Moreover, $\log \Phi_X$ is continuous and vanishes at $t = 0$. Now assume for some $r \geq 1$, $E|X|^r$ exists, we then have

$$\Phi_X(t) = \sum_{j=0}^{r} \mu_j' \frac{(it)^j}{r!} + o(|t|^r), \tag{2.2}$$

as $t \to 0$. We can then use the Maclaurin series $\log(1 + z) = z - z^2/2 + z^3/3 - \cdots$ to obtain

$$\log \Phi_X(t) = \sum_{j=1}^{r} \kappa_j \frac{(it)^j}{r!} + o(|t|^r), \tag{2.3}$$

as $t \to 0$, where the coefficients, κ_j's, are the cumulants by definition. It is now evident that the cumulant of order r exists if the moments of order r and lower exist.

6

Joint cumulants that involve two or more random variables are also of interest. Consider random variables $X_1, \ldots, X_r$, such that $E|X_j|^n < \infty$, $j = 1, \ldots, r$, for a positive integer n. Then all mixed absolute moments and cumulants of $X_1, \ldots, X_r$, up to order n exist. Let $\Phi_{X_1 \cdots X_r}(t_1, \ldots, t_r) = E\left\{ \exp\left[i \sum_{j=1}^{r} t_j X_j \right] \right\}$ be the joint characteristic function of $X_1, \ldots, X_r$, then joint cumulants of $X_1, \ldots, X_r$, $\mathrm{cum}(X_1, \ldots, X_r)$, is given by the coefficient of the $(i)^r t_1 \cdots t_r / r!$ term in the Taylor series expansion of $\log \Phi_{X_1 \cdots X_r}(\cdot)$ about the origin.

According to the above definition, the first four joint cumulants are

$$\mathrm{cum}(X_1) = EX_1,$$

$$\mathrm{cum}(X_1, X_2) = E\left[(X_1 - EX_1)(X_2 - EX_2) \right] = \mathrm{Cov}(X_1, X_2),$$

$$\mathrm{cum}(X_1, X_2, X_3) = E\left[(X_1 - EX_1)(X_2 - EX_2)(X_3 - EX_3) \right],$$

and

$$\mathrm{cum}(X_1, X_2, X_3, X_4) = E\left[(X_1 - EX_1)(X_2 - EX_2)(X_3 - EX_3)(X_4 - EX_4) \right]$$
$$- \mathrm{Cov}(X_1, X_2)\mathrm{Cov}(X_3, X_4) - \mathrm{Cov}(X_1, X_3)\mathrm{Cov}(X_2, X_4)$$
$$- \mathrm{Cov}(X_1, X_4)\mathrm{Cov}(X_2, X_3).$$

In general, the rth order joint cumulant, $\mathrm{cum}(X_1, \ldots, X_r)$, is given by

$$\mathrm{cum}(X_1, \ldots, X_r) = \sum (-1)^{p-1}(p-1)! \left(E \prod_{j \in \nu_1} X_j \right) \cdots \left(E \prod_{j \in \nu_p} X_j \right), \qquad (2.4)$$

where the summation extends over all partitions $\nu_1, \ldots, \nu_p$, $p = 1, \ldots, r$, of the set $\{1, \ldots, r\}$.

Representing joint moments in terms of joint cumulants are also useful. We

record those that will be used.

$$E[X_1 X_2] = \text{cum}(X_1, X_2) + \text{cum}(X_1)\text{cum}(X_2),$$

$$E[X_1 X_2 X_3] = \text{cum}(X_1, X_2, X_3) + \text{cum}(X_1)\text{cum}(X_2, X_3) + \text{cum}(X_2)\text{cum}(X_1, X_3)$$

$$+ \text{cum}(X_3)\text{cum}(X_1, X_2) + \text{cum}(X_1)\text{cum}(X_2)\text{cum}(X_3),$$

$$E[X_1 X_2 X_3 X_4] = \text{cum}(X_1, X_2, X_3, X_4) + \text{cum}(X_1)\text{cum}(X_2, X_3, X_4)$$

$$+ \text{cum}(X_2)\text{cum}(X_1, X_3, X_4) + \text{cum}(X_3)\text{cum}(X_1, X_2, X_4)$$

$$+ \text{cum}(X_4)\text{cum}(X_1, X_2, X_3) + \text{cum}(X_1, X_2)\text{cum}(X_3, X_4)$$

$$+ \text{cum}(X_1, X_3)\text{cum}(X_2, X_4) + \text{cum}(X_1, X_4)\text{cum}(X_2, X_3)$$

$$+ \text{cum}(X_1, X_2)\text{cum}(X_3)\text{cum}(X_4) + \text{cum}(X_1, X_3)\text{cum}(X_2)\text{cum}(X_4)$$

$$+ \text{cum}(X_1, X_4)\text{cum}(X_2)\text{cum}(X_3) + \text{cum}(X_3, X_4)\text{cum}(X_1)\text{cum}(X_2)$$

$$+ \text{cum}(X_2, X_4)\text{cum}(X_1)\text{cum}(X_3) + \text{cum}(X_2, X_3)\text{cum}(X_1)\text{cum}(X_4)$$

$$+ \text{cum}(X_1)\text{cum}(X_2)\text{cum}(X_3)\text{cum}(X_4).$$

For convenience, if i, j, and k are positive integers, for random variables X. Y, and Z, we define

$$\kappa_i(X) = \text{cum}(\underbrace{X, \ldots, X}_{i \text{ terms}}), \tag{2.5}$$

$$\kappa_{ij}(X, Y) = \text{cum}(\underbrace{X, \ldots, X}_{i \text{ terms}}, \underbrace{Y, \ldots, Y}_{j \text{ terms}}), \tag{2.6}$$

and

$$\kappa_{ijk}(X, Y, Z) = \text{cum}(\underbrace{X, \ldots, X}_{i \text{ terms}}, \underbrace{Y, \ldots, Y}_{j \text{ terms}}, \underbrace{Z, \ldots, Z}_{k \text{ terms}}). \tag{2.7}$$

We record here some useful properties of cumulants; see Brillinger [6], p. 19. Note that in all cases p. r. and s are positive integers.

8

(1) $\mathrm{cum}(a_1 X_1, \ldots, a_r X_r) = a_1 \cdots a_r \mathrm{cum}(X_1, \ldots, X_r)$ for constants $a_1, \ldots, a_r$.

(2) $\mathrm{cum}(X_1, \ldots, X_r)$ is symmetric in its arguments so that $\mathrm{cum}(X_1, \ldots, X_r) = \mathrm{cum}(X_{\pi_1}, \ldots, X_{\pi_r})$ for any permutation $(\pi_1, \ldots, \pi_r)$ of $(1, \ldots, r)$.

(3) $\mathrm{cum}(X_1, \ldots, X_r) = 0$ if any nontrivial proper subset of X_i's are independent of the remaining X_i's. To see this, we may suppose that $(X_1, \ldots, X_p)$ and $(X_{p+1}, \ldots, X_r)$ are independent, then

$$\log\left\{E \exp\left[i \sum_{j=1}^{r} t_j X_j\right]\right\} = \log\left\{E \exp\left[i \sum_{j=1}^{p} t_j X_j\right]\right\} + \log\left\{E \exp\left[i \sum_{j=p+1}^{r} t_j X_j\right]\right\}.$$

There will be no $(i)^r t_1 \cdots t_r$ term in the Taylor series expansion on the right hand side; the desired result now follows from equation (2.4), the definition of joint cumulants.

(4) $\mathrm{cum}(X_1 + Y, X_2, \ldots, X_r) = \mathrm{cum}(X_1, X_2, \ldots, X_r) + \mathrm{cum}(Y, X_2, \ldots, X_r)$.

(5) $\mathrm{cum}(X_1 + a_1, \ldots, X_r + a_r) = \mathrm{cum}(X_1, \ldots, X_r)$ when $r \geq 2$ and a_i's are constants. It is sufficient to note from (4), that $\mathrm{cum}(X_1 + a_1, X_2, \ldots, X_r) = \mathrm{cum}(X_1, X_2, \ldots, X_r) + \mathrm{cum}(a_1, X_2, \ldots, X_r)$, but the last term is zero from (3). The result now follows from induction.

(6) For a random variable X the cumulants $\{\kappa_r(X)\}$ and moments $\{\mu'_r(X)\}$ satisfy the following relationship.

$$\kappa_r = \sum_{s=0}^{r} \sum (-1)^{m-1} \frac{(m-1)! \, r!}{i_1! \cdots i_s!} \left(\frac{\mu'_{j_1}}{j_1!}\right)^{i_1} \cdots \left(\frac{\mu'_{j_s}}{j_s!}\right)^{i_s}, \qquad (2.8)$$

$$\mu'_r = \sum_{s=0}^{r} \sum \frac{r!}{i_1! \cdots i_s!} \left(\frac{\kappa_{j_1}}{j_1!}\right)^{i_1} \cdots \left(\frac{\kappa_{j_s}}{j_s!}\right)^{i_s}, \qquad (2.9)$$

9

where both second summations extend over all nonnegative integers i_s's and m's, and positive integers j_s's, such that $\sum_{t=1}^{s} i_t = m$ and $\sum_{t=1}^{s} i_t j_t = r$ (Lukacs [20], p. 27).

(7) For a random variable X the cumulants $\{\kappa_r(X)\}$ and central moments $\{\mu_r(X)\}$ satisfy the following relationship.

$$\kappa_r = \sum_{s=0}^{r} \sum (-1)^{m-1} \frac{(m-1)! \, r!}{i_1! \cdots i_s!} \left(\frac{\mu_{j_1}}{j_1!}\right)^{i_1} \cdots \left(\frac{\mu_{j_s}}{j_s!}\right)^{i_s}, \qquad (2.10)$$

$$\mu_r = \sum_{s=0}^{r} \sum \frac{r!}{i_1! \cdots i_s!} \left(\frac{\kappa_{j_1}}{j_1!}\right)^{i_1} \cdots \left(\frac{\kappa_{j_s}}{j_s!}\right)^{i_s}, \qquad (2.11)$$

where both second summations extend over all nonnegative integers i_s's and m's, and integers j_s's, $j_s > 1$ for each s, such that $\sum_{t=1}^{s} i_t = m$ and $\sum_{t=1}^{s} i_t j_t = r$. To see this, let $X' = X - EX$. Notice that $\kappa_1(X') = \mu_1'(X') = \mu_1(X') = 0$; $\kappa_i(X') = \kappa_i(X)$ and $\mu_i'(X') = \mu_i(X') = \mu_i(X)$, for each $i \geq 2$. Applying (6) to $\kappa_i(X')$'s and $\mu_i'(X')$'s yields the desired result.

(8) Joint moments can be representing by sums of products of joint cumulants.

$$EX_1 \cdots X_r = \sum \kappa_{\nu_1} \cdots \kappa_{\nu_p}, \qquad (2.12)$$

where the summation extends over all partitions $\nu_1, \ldots, \nu_p$, $p = 1, \ldots, r$, of the set $\{1, \ldots, r\}$, and

$$\kappa_{\nu_s} = \mathrm{cum}(X_{\alpha_1}, \ldots, X_{\alpha_m}), \qquad (2.13)$$

where the α_j's are the elements of ν_s (Rosenblatt [22], p. 34).

2.2. Formal Edgeworth Expansion

We now discuss some results for asymptotic expansions of distribution functions. In order to motivate the idea of the expansions, we proceed formally as follows.

Suppose F_X, Φ_X, $\{\kappa_r(X) : r \geq 1\}$, and F_Y, Φ_Y, $\{\kappa_r(Y) : r \geq 1\}$ are distribution functions, characteristic functions, and cumulants of random variables X and Y, respectively. Since $\Phi_X(t) = \exp\left\{ \sum_{r=1}^{\infty} \kappa_r(X)(it)^r/r! \right\}$ and $\Phi_Y(t) = \exp\left\{ \sum_{r=1}^{\infty} \kappa_r(Y)(it)^r/r! \right\}$, we have

$$\Phi_X(t) = \exp\left\{ \sum_{r=1}^{\infty} [\kappa_r(X) - \kappa_r(Y)] \frac{(it)^r}{r!} \right\} \Phi_Y(t). \qquad (2.14)$$

If the characteristic function Φ_Y is absolutely integrable over $(-\infty, \infty)$, then F_Y is absolutely continuous and

$$f_Y(y) \equiv F_Y'(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\{-ity\} \Phi_Y(t) dt. \qquad (2.15)$$

The density f_Y is bounded and continuous (Chung [8], p. 155). It follows that, subject to existence,

$$D^r f_Y(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\{-ity\} (-it)^r \Phi_Y(t) dt. \qquad (2.16)$$

where $D \equiv d/dy$ denotes the differential operator. Then, formally, the characteristic function of the distribution function $D^r F_Y$ will be $(-it)^r \Phi_Y(t)$. By this observation and from equation (2.14), the uniqueness property of the Fourier transform

11

then yields

$$F_X(x) = \exp\left\{ \sum_{r=1}^{\infty} [\kappa_r(X) - \kappa_r(Y)] \frac{(-D)^r}{r!} \right\} F_Y(x), \qquad (2.17)$$

and similarly

$$f_X(x) = \exp\left\{ \sum_{r=1}^{\infty} [\kappa_r(X) - \kappa_r(Y)] \frac{(-D)^r}{r!} \right\} f_Y(x). \qquad (2.18)$$

The special case where Y is a normal random variable is most important. Let Y be $N(\mu, \sigma^2)$ and set

$$\alpha(y) \equiv (1/\sqrt{2\pi}) \exp\{-y^2/2\}. \qquad (2.19)$$

Then Y has density function $f_Y(y) \equiv \sigma^{-1}\alpha((y-\mu)/\sigma)$, characteristic function $\Phi_Y(t) = \exp\{i\mu t - \sigma^2 t^2/2\}$, and cumulants $\kappa_1(Y) = \mu$, $\kappa_2(Y) = \sigma^2$, and $\kappa_r(Y) = 0$. for each $r \geq 3$. Equation (2.18) becomes

$$f_X(x) = \exp\left\{ -\frac{\kappa_1(X) - \mu}{1!}D + \frac{\kappa_2(X) - \sigma^2}{2!}D^2 \right.$$
$$\left. -\frac{\kappa_3(X)}{3!}D^3 + \frac{\kappa_4(X)}{4!}D^4 - \cdots \right\} f_Y(x). \qquad (2.20)$$

This is the well known Edgeworth expansion of Type A (Kendall and Stuart [16], p. 170).

Another form of the Type A series is also of interest. First note that the various derivatives of f_Y in the expansion can be expressed in the form of Chebyshev-Hermite polynomials. Specifically, we have (Kendall and Stuart [16], p. 167)

$$(-D)^r \alpha(x) = H_r(x)\alpha(x), \qquad (2.21)$$

12

where $\{H_r(x) : i \geq 0\}$ are the Chebyshev-Hermite polynomials given by:

$$H_0(x) = 1,$$

$$H_1(x) = x,$$

$$H_2(x) = x^2 - 1,$$

$$H_3(x) = x^3 - 3x,$$

$$H_4(x) = x^4 - 6x^2 + 3,$$

and, in general, the rth Chebyshev-Hermite polynomial has the form

$$H_r(x) = \sum_{j=0}^{\lceil r/2 \rceil} x^{r-2j}(-1)^j \frac{r!}{(r-2j)!j!2^j}, \tag{2.22}$$

where $\lceil x \rceil$ denotes the smallest integer which is greater than or equal to x. Since $f_Y(x) = \alpha((x - \mu)/\sigma)$, from equation (2.21),

$$(-D)^r f_Y(x) = \sigma^{-r} H_r((x - \mu)/\sigma) f_Y(x). \tag{2.23}$$

From equations (2.18) and (2.23) we can see that f_X can now be formally expanded in the products of f_Y and the Chebyshev-Hermite polynomials.

Up to this point an underlying assumption is that the functions f_X and F_X possess convergent Type A series. This is not always the case. For a discussion of the convergence properties see Kendall and Stuart [16], pp. 173–174, and Cramér [10], p. 223.

For practical applications, however, it is usually of little value to know the convergence properties of the expansions. What we are really interested in is whether a small number of terms would suffice to produce good approximations of the functions f_X and F_X. If this is the case, we would not be too concerned about the

13

convergence properties. On the other hand, if the series actually converges but a satisfactory approximation can only be obtained after a large number of terms have been calculated, then this Type A series would be of very little use.

One important application of the Type A series is when the cumulants of the random variables have a special structure. For example, suppose we are given a sequence of random variables $\{S_n : n \geq 1\}$, with $ES_n = 0$, $n = 1, 2, \ldots$, and for each $i \geq 2$, the ith cumulant of S_n,

$$\kappa_i(S_n) = O(n), \tag{2.24}$$

as $n \to \infty$. We are interested in the expansion of the distribution function of the normalized random variable

$$Z_n \equiv S_n / \kappa_2^{1/2}(S_n), \tag{2.25}$$

as $n \to \infty$.

We use this notation: $\kappa_i \equiv \kappa_i(S_n)$ and $l_1 \equiv \kappa_1(Z_n) = \kappa_1 / \kappa_2^{1/2}$, $l_2 \equiv \kappa_2(Z_n) - 1 = 0$, and $l_i \equiv \kappa_i(Z_n) = \kappa_i / \kappa_2^{i/2}$, for each $i \geq 3$. From equations (2.24) and (2.25), it follows that, for each $i \geq 3$,

$$l_i = O(n^{1-i/2}). \tag{2.26}$$

Since $\kappa_1(Z_n) = 0$ and $\kappa_2(Z_n) = 1$ we will choose $\mu = 0$ and $\sigma^2 = 1$ in equation (2.18), the Edgeworth expansion of Type A, and obtain

$$f_{Z_n}(x) = \exp\left\{ -\frac{l_3}{6} D^3 + \frac{l_4}{24} D^4 - \frac{l_5}{120} D^5 + \cdots \right\} \alpha(x). \tag{2.27}$$

Using

$$\exp\{n^{-1/2} d_1 + n^{-1} d_2 + n^{-3/2} d_3 + \cdots\}$$

14

$$=1 + n^{-1/2}\Big[d_1\Big] + n^{-1}\Big[\frac{1}{2}d_1^2 + d_2\Big] + n^{-3/2}\Big[\frac{1}{6}d_1^3 + d_1 d_2 + d_3\Big] + \cdots, \quad (2.28)$$

and equation (2.26), we can, at least formally, approximate f_{Z_n} by the derivatives of $\alpha(\cdot)$ up to any order of $n^{-1/2}$. For our purpose, we neglect terms of order $o(n^{-3/2})$, then we have, as $n \to \infty$

$$f_{Z_n}(x) = \exp\Big\{ -\frac{l_3}{6}D^3 + \frac{l_4}{24}D^4 - \frac{l_5}{120}D^5\Big\}\alpha(x) + o(n^{-3/2})$$

$$= \Big\{1 - \frac{l_3}{6}D^3 + \Big[\frac{l_4}{24}D^4 + \frac{l_3^2}{72}D^6\Big]$$

$$- \Big[\frac{l_3^3}{1296}D^9 + \frac{l_3 l_4}{144}D^7 + \frac{l_5}{120}D^5\Big]\Big\}\alpha(x) + o(n^{-3/2})$$

$$= \alpha(x) + \frac{l_3}{6}H_3(x)\alpha(x) + \Big[\frac{l_4}{24}H_4(x) + \frac{l_3^2}{72}H_6(x)\Big]\alpha(x)$$

$$+ \Big[\frac{l_3^3}{1296}H_9(x) + \frac{l_3 l_4}{144}H_7(x) + \frac{l_5}{120}H_5(x)\Big]\alpha(x) + o(n^{-3/2}). \quad (2.29)$$

If we apply similar argument to F_{Z_n}, we then have (Cramér [10]), as $n \to \infty$

$$F_{Z_n}(x) = \Phi(x) - \frac{l_3}{6}H_2(x)\alpha(x) - \Big[\frac{l_4}{24}H_3(x) + \frac{l_3^2}{72}H_5(x)\Big]\alpha(x)$$

$$- \Big[\frac{l_3^3}{1296}H_8(x) + \frac{l_3 l_4}{144}H_6(x) + \frac{l_5}{120}H_4(x)\Big]\alpha(x) + o(n^{-3/2}). \quad (2.30)$$

Note that if we only include the first term on the right hand side this is the usual central limit theorem while all the remaining terms, with coefficients of smaller order of n, represent the error terms in the approximation of the central limit theorem.

In practical applications, we may not know the exact value for $\kappa_2(S_n)$, which is the the variance of S_n. Suppose that $\widehat{\kappa_2(S_n)}$ is an estimator for $\kappa_2(S_n)$ then a

15

formal Edgeworth expansion of $\widehat{Z_n} \equiv S_n / \kappa_2(\widehat{S_n})^{1/2}$ will be of interest.

If $\widehat{l_i} \equiv \kappa_i(\widehat{Z_n})$, for each $i \geq 1$, satisfies some asymptotic properties, then a formal Edgeworth expansion can also be obtained. For example, suppose $\widehat{l_1} = O(n^{-1/2})$, $\widehat{l_2} = O(n^{-1})$, and for each $i \geq 3$, $\widehat{l_i} = O(n^{1-i/2})$. Then it can be shown that, as $n \to \infty$ (Kendall and Stuart [16], p. 176)

$$
\begin{aligned}
F_{\widehat{Z_n}}(x) = \Phi(x) &- \left[\widehat{l_1} + \frac{\widehat{l_3}}{6} H_2(x)\right]\alpha(x) \\
&- \left[\frac{\widehat{l_1}^2}{2}H_1(x) + \frac{\widehat{l_2}}{2}H_1(x)\right. \\
&\left. + \frac{\widehat{l_1}\widehat{l_3}}{6}H_3(x) + \frac{\widehat{l_4}}{24}H_3(x) + \frac{\widehat{l_3}^2}{72}H_5(x)\right]\alpha(x) + o(n^{-1}). \quad (2.31)
\end{aligned}
$$

2.3. Cornish-Fisher Expansions

Suppose that we have a sequence of random variables $\{Y_n : n \geq 1\}$ which are asymptotically normal in the sense that there exist sequences $\{\mu_n : n \geq 1\}$ and $\{\sigma_n : n \geq 1\}$ such that as $n \to \infty$,

$$
\frac{Y_n - \mu_n}{\sigma_n} \Rightarrow N(0, 1). \quad (2.32)
$$

For large n, it is often possible to approximate the distribution of $(Y_n - \mu_n)/\sigma_n$ by a normal, but for small to moderate n this may not be a good approximation. Under some circumstance we are able to use a polynomial-variate transformation such as

$$
\xi_n = b_{n,0} + b_{n,1}\left(\frac{Y_n - \mu_n}{\sigma_n}\right) + b_{n,2}\left(\frac{Y_n - \mu_n}{\sigma_n}\right)^2 + \cdots, \quad (2.33)
$$

16

where $b_{n,i}$'s are of order $n^{1/2}$ or smaller. By choosing $b_{n,i}$'s appropriately, we may make the distribution of ξ_n much closer to normal than that of $(Y_n - \mu_n)/\sigma_n$.

It turns out that this problem is embedded in the following larger and more natural question. Suppose ξ is a $N(0,1)$ random variable, Z_n has distribution function F_{Z_n} and and its cumulants satisfy equation (2.26). If x and ξ are corresponding quantiles of F_{Z_n} and Φ respectively, namely,

$$F_{Z_n}(x) = \Phi(\xi), \tag{2.34}$$

solving, at least formally, ξ in terms of x and x in terms of ξ will be of interest.

We sketch how to solve equation (2.34) when F_{Z_n} can only be estimated. The basic idea is to approximate F_{Z_n} by its Edgeworth expansion. Let z be the difference between distribution functions F_{Z_n} and Φ, then as $n \to \infty$

$$z \equiv F_{Z_n}(x) - \Phi(x)$$

$$= -[\frac{l_3}{6}H_2(x)]\alpha(x) - [\frac{l_4}{24}H_3(x) + \frac{l_3^2}{72}H_5(x)]\alpha(x) - \cdots. \tag{2.35}$$

If we neglect terms of $o(n^{-1})$, then

$$z = -[\frac{l_3}{6}H_2(x)]\alpha(x) - [\frac{l_4}{24}H_3(x) + \frac{l_3^2}{72}H_5(x)]\alpha(x) + o(n^{-1}). \tag{2.36}$$

Then from equations (2.34), (2.36), and a Taylor expansion,

$$\xi = \Phi^{-1}(F_{Z_n}(x))$$

$$= \Phi^{-1}(\Phi(x) + z)$$

$$= \Phi^{-1}(\Phi(x)) + z[\Phi^{-1}]'(\Phi(x)) + \frac{z^2}{2}[\Phi^{-1}]''(\Phi(x))$$

17

$$+ \frac{z^3}{6} [\Phi^{-1}]'''(\Phi(x)) + \cdots . \tag{2.37}$$

To evaluate $[\Phi^{-1}]'(\Phi(\cdot))$, $[\Phi^{-1}]''(\Phi(\cdot))$, and $[\Phi^{-1}]'''(\Phi(\cdot))$, we need the following observation. Suppose $\varphi(\cdot)$ is a monotone function and $\psi(\cdot) \equiv \varphi^{-1}(\cdot)$ is its inverse function. If the third derivatives of both $\varphi(\cdot)$ and $\psi(\cdot)$ exist, then it can be shown that

$$[\psi]'(v) = \frac{1}{\varphi'(\psi(v))}, \tag{2.38}$$

$$[\psi]''(v) = -\frac{\varphi''(\psi(v))}{[\varphi'(\psi(v))]^3}, \tag{2.39}$$

$$[\psi]'''(v) = 3\frac{[\varphi''(\psi(v))]^2}{[\varphi'(\psi(v))]^5} - \frac{\varphi'''(\psi(v))}{[\varphi'(\psi(v))]^4}. \tag{2.40}$$

From equations (2.21), (2.38), (2.39), and (2.40), it follows that

$$[\Phi^{-1}]'(\Phi(x)) = 1/\alpha(x), \tag{2.41}$$

$$[\Phi^{-1}]''(\Phi(x)) = H_1(x)/[\alpha(x)]^2, \tag{2.42}$$

$$[\Phi^{-1}]'''(\Phi(x)) = [3H_1^2(x) - H_2(x)]/[\alpha(x)]^3. \tag{2.43}$$

Finally, observe that z is a polynomial of $n^{-1/2}$, again, we can obtain a formal approximation of (2.34) up to any order of $n^{-1/2}$. Substitute the approximation (2.36) into equation (2.37) and truncate those terms of order $o(n^{-1})$. We obtain

$$\xi = x - \left[\frac{l_3}{6}(x^2 - 1)\right] + \left[\frac{l_3^2}{36}(4x^3 - 7x) - \frac{l_4}{24}(x^3 - 3x)\right] + o(n^{-1}). \tag{2.44}$$

Notice that the three terms on the right hand side are of orders $O(1)$, $O(n^{-1/2})$, and $O(n^{-1})$, respectively.

18

In practice, it is sometimes more convenient to express x in terms of ξ. This may be done by a technique introduced by Cornish and Fisher [9]. First, let η represents the difference between quantile points x and ξ,

$$\eta(x) \equiv x - \xi, \tag{2.45}$$

From equation (2.44), it follows that

$$\eta(\xi) = \frac{l_3}{6}(\xi^2 - 1) + \frac{l_4}{24}(\xi^3 - 3\xi) - \frac{l_3^2}{36}(4\xi^3 - 7\xi) + o(n^{-1}) = O(n^{-1/2}), \tag{2.46}$$

$$\eta'(\xi) = \frac{l_3}{3}(\xi) + \frac{l_4}{8}(\xi^2 - 1) - \frac{l_3^2}{36}(12\xi^2 - 7) + o(n^{-1}) = O(n^{-1/2}), \tag{2.47}$$

$$\eta''(\xi) = \frac{l_3}{3} + \frac{l_4}{4}(\xi) - \frac{l_3^2}{3} + o(n^{-1}) = O(n^{-1/2}). \tag{2.48}$$

Then observe that by a Taylor expansion,

$$x - \xi = \eta(\xi + x - \xi)$$

$$= \eta(\xi) + (x - \xi)\eta'(\xi) + \frac{1}{2}(x - \xi)^2\eta''(\xi) + \cdots$$

$$= \eta(\xi) + \left\{ \eta(\xi) + (x - \xi)\eta'(\xi) + \frac{1}{2}(x - \xi)^2\eta''(\xi) + \cdots \right\}\eta'(\xi)$$

$$+ \frac{1}{2}\left\{ \eta(\xi) + (x - \xi)\eta'(\xi) + \frac{1}{2}(x - \xi)^2\eta''(\xi) + \cdots \right\}^2\eta''(\xi)$$

$$+ \cdots, \tag{2.49}$$

where the last equality is obtained by substituting the value of $x - \xi$ on the right hand side of the second equality. Using this technique again and again yields, at least formally,

$$x - \xi = \eta(\xi) + \eta(\xi)\eta'(\xi) + \left\{ \eta(\xi)[\eta'(\xi)]^2 + \frac{1}{2}\eta^2(\xi)\eta''(\xi) \right\}$$

$$+ \left\{ \eta(\xi)[\eta'(\xi)]^3 + \frac{3}{2}\eta^2(\xi)\eta'(\xi)\eta''(\xi) + \frac{1}{6}\eta^3(\xi)\eta'''(\xi) \right\} + \cdots. \tag{2.50}$$

19

Finally, substitute approximations (2.46), (2.47) and (2.48) into (2.50) and neglect terms of $o(n^{-1})$, we then have

$$x = \xi + \frac{l_3}{6}(\xi^2 - 1) + \left\{ \frac{l_4}{24}(\xi^3 - 3\xi) - \frac{l_3^2}{36}(2\xi^3 - 5\xi) \right\} + o(n^{-1}). \qquad (2.51)$$

This is the required representation of x in ξ. Notice that the three terms on the right hand side are of orders $O(1)$, $O(n^{-1/2})$, and $O(n^{-1})$, respectively.

Both equations (2.44) and (2.51) are approximated to order $O(n^{-1})$. Cornish and Fisher [9] also give higher-order expansions and tables to facilitate their use.

The above Cornish-Fisher expansions (2.44) and (2.51) correspond to the Edgeworth expansion (2.30), namely, as $n \to \infty$

$$F_{Z_n}(x) = \Phi(x) - \frac{l_3}{6}H_2(x)\alpha(x) - \left[\frac{l_4}{24}H_3(x) + \frac{l_3^2}{72}H_5(x) \right]\alpha(x) + o(n^{-1}). \qquad (2.52)$$

Recall that the Edgeworth expansion of the form of equation (2.31), i.e.

$$F_{\widehat{Z}_n}(x) = \Phi(x) - \left[\widehat{l_1} + \frac{\widehat{l_3}}{6}H_2(x) \right]\alpha(x)$$

$$- \left[\frac{\widehat{l_1}^2}{2}H_1(x) + \frac{\widehat{l_2}}{2}H_1(x) + \frac{\widehat{l_1}\widehat{l_3}}{6}H_3(x) + \frac{\widehat{l_4}}{24}H_3(x) + \frac{\widehat{l_3}^2}{72}H_5(x) \right]\alpha(x)$$

$$+ o(n^{-1}), \qquad (2.53)$$

is useful when we do not have the exact value of the variance of a random variable. The corresponding Cornish-Fisher expansions can be obtained by applying the same procedures. We list the results as follows.

$$\xi = x - \left[\widehat{l_1} + \frac{\widehat{l_3}}{6}(x^2 - 1) \right] - \left[\frac{\widehat{l_2}}{2}x - \frac{\widehat{l_1}\widehat{l_3}}{3}(x) \right]$$

20

$$+ \left[\frac{\widehat{l_3}^2}{36}(4x^3 - 7x) - \frac{\widehat{l_4}}{24}(x^3 - 3x) \right] + o(n^{-1}) \qquad (2.54)$$

$$x = \xi + \left[\widehat{l_1} + \frac{\widehat{l_3}}{6}(\xi^2 - 1) \right] + \frac{\widehat{l_2}}{2}\xi$$

$$+ \left[\frac{\widehat{l_4}}{24}(\xi^3 - 3\xi) - \frac{\widehat{l_3}^2}{36}(2\xi^3 - 5\xi) \right] + o(n^{-1}). \qquad (2.55)$$

An application of the formal Cornish-Fisher expansions (2.44), (2.51), (2.54), and (2.55) is as follows. It is quite often in the simulation study that we only know a distribution function F_{Z_n} is asymptotically standard normal and we are interested in its δ quantile point x_δ. Traditionally we are used to let the corresponding quantile point ξ_δ of Φ as a substitute for x_δ, namely,

$$\widehat{x_\delta} = \xi_\delta. \qquad (2.56)$$

But according to the Cornish-Fisher expansions (2.54) and (2.55), if we either know or have a method to estimate the normalized cumulants l_i's of Z_n, then the corresponding (estimated) adjusted quantile points

$$h_1(\xi_\delta) = \xi_\delta + \widehat{l_1} + \frac{\widehat{l_3}}{6}(\xi_\delta^2 - 1) \qquad (2.57)$$

and

$$h_2(\xi_\delta) = \xi_\delta + \widehat{l_1} + \frac{\widehat{l_3}}{6}(\xi_\delta^2 - 1) + \frac{\widehat{l_2}}{2}\xi_\delta + \frac{\widehat{l_4}}{24}(\xi_\delta^3 - 3\xi_\delta) - \frac{\widehat{l_3}^2}{36}(2\xi_\delta^3 - 5\xi_\delta) \qquad (2.58)$$

will be both asymptotically better, as estimators for x_δ, than ξ_δ.

2.4. Uniqueness Properties of Cornish-Fisher Expansions

There are two interesting interpretations of the Cornish-Fisher expansions. We state them below but leave their proofs to the Appendix.

Proposition 2.1. *Suppose ξ and X_n are random variables, where ξ is $N(0,1)$ and X_n satisfies $\kappa_1(X_n) = l_{n,1} = O(n^{-1/2})$, $\kappa_2(X_n) = 1 + l_{n,2} = 1 + O(n^{-1})$, $\kappa_3(X_n) = l_{n,3} = O(n^{-1/2})$, $\kappa_4(X_n) = l_{n,4} = O(n^{-1})$, and $\kappa_i(X_n) = O(n^{-1})$, $i \geq 5$. Then.*

(1) there is a polynomial of degree two, $g_1(X_n) = X_n + a_{n,0} + a_{n,1}X_n + a_{n,2}X_n^2$, such that $a_{n,i} = O(n^{-1/2})$, $i = 0, 1, 2$, and $|\kappa_i(\xi) - \kappa_i(g_1(X_n))| = O(n^{-1})$ for each i, $1 \leq i \leq 3$. Moreover, the coefficients $a_{n,i}$'s are unique up to $O(n^{-1/2})$. Specifically, $a_{n,0} = -l_{n,1} + (1/6)l_{n,3} + o(n^{-1/2})$, $a_{n,1} = o(n^{-1/2})$, and $a_{n,2} = -(1/6)l_{n,3} + o(n^{-1/2})$.

(2) If in addition, $\kappa_i(X_n) = o(n^{-1})$, $i \geq 5$, then there is a polynomial of degree three, $g_2(X_n) = X_n + a_{n,0} + a_{n,1}X_n + a_{n,2}X_n^2 + a_{n,3}X_n^3$, such that $a_{n,i} = O(n^{-1/2})$ for each i, $1 \leq i \leq 4$, and $|\kappa_i(\xi) - \kappa_i(g_2(X_n))| = o(n^{-1})$ for each i, $1 \leq i \leq 4$. Moreover, the coefficients $a_{n,i}$'s are unique up to $O(n^{-1})$. Specifically, $a_{n,0} = -l_{n,1} + (1/6)l_{n,3} + o(n^{-1})$, $a_{n,1} = -(1/2)l_{n,2} + (1/3)l_{n,1}l_{n,3} + (1/8)l_{n,4} - (7/36)l_{n,3}^2 + o(n^{-1})$. $a_{n,2} = -(1/6)l_{n,3} + o(n^{-1})$, and $a_{n,3} = -(1/24)l_{n,4} + (1/9)l_{n,3}^2 + o(n^{-1})$.

Proposition 2.2. *Suppose ξ and X_n are random variables, where ξ is $N(0,1)$ and X_n satisfies $\kappa_1(X_n) = l_{n,1} = O(n^{-1/2})$, $\kappa_2(X_n) = 1 + l_{n,2} = 1 + O(n^{-1})$. $\kappa_3(X_n) = l_{n,3} = O(n^{-1/2})$. Then,*

(1) there is a polynomial of degree two, $h_1(\xi) = \xi + a_{n,0} + a_{n,1}\xi + a_{n,2}\xi^2$. such that $a_{n,i} = O(n^{-1/2})$, $i = 0, 1, 2$, and $|\kappa_i(X_n) - \kappa_i(h_1(\xi))| = O(n^{-1})$ for each i.

$1 \leq i \leq 3$. *Moreover, the coefficients $a_{n,i}$'s are unique up to $O(n^{-1/2})$. Specifically,*
$a_{n,0} = l_{n,1} - (1/6)l_{n,3} + o(n^{-1/2})$, $a_{n,1} = o(n^{-1/2})$, *and* $a_{n,2} = (1/6)l_{n,3} + o(n^{-1/2})$.
(2) If in addition $\kappa_4(X_n) = l_{n,4} = O(n^{-1})$, then there is a polynomial of degree
three, $h_2(\xi) = \xi + a_{n,0} + a_{n,1}\xi + a_{n,2}\xi^2 + a_{n,3}\xi^3$, such taht $a_{n,i} = O(n^{-1/2})$ for each
i, $1 \leq i \leq 4$, and $|\kappa_i(X_n) - \kappa_i(h_2(\xi))| = o(n^{-1})$ for each i, $1 \leq i \leq 4$. Moreover,
the coefficients $a_{n,i}$'s are unique up to $O(n^{-1})$. Specifically, $a_{n,0} = l_{n,1} - (1/6)l_{n,3} +$
$o(n^{-1})$, $a_{n,1} = (1/2)l_{n,2} - (1/8)l_{n,4} + (5/36)l_{n,3}^2 + o(n^{-1})$, $a_{n,2} = (1/6)l_{n,3} + o(n^{-1})$,
and $a_{n,3} = (1/24)l_{n,4} - (1/18)l_{n,3}^2 + o(n^{-1})$.

In Proposition 2.1 we can actually obtain a slightly better result. For example.
in part (1) we can calculate $a_{n,i}$'s so that they will be unique to order $O(n^{-1})$.
However, according to our assumption of the cumulants of X_n, the neglected terms
in the formal Cornish-Fisher expansion of X_n are of order $O(n^{-1})$. Thus calculating
a_i's to a lower order will not produce a better approximation, at least orderwise.
For this reason we do not present that result. A similar result and observation also
apply to part (2).

One significant point of Propositions 2.1 and 2.2 is that both the transforma-
tions from a standard normal random variable to a statistic and the transforma-
tion from a statistic to a standard normal random variable are unique. Moreover.
the second and third order polynomial transformations coincide with the first two
Cornish-Fisher expansions; see equations (2.54) and (2.55).

As a consequence of an Edgeworth expansion, Cornish-Fisher expansions can
transform a statistic so that its distribution is closer to a standard normal random
variable. Thus it can be used to generate an asymptotically better confidence
interval. Unfortunately. for many statistics associated with real world processes. it

is extremely difficult to prove the existence of a rigorous Edgeworth expansion. The two propositions above, in a sense, provide us an alternative approach for obtaining "formal" Cornish-Fisher expansions for these situations.

2.5. Cumulants of Stationary Processes

In this section, we will study some properties of the cumulants of stationary processes.

Suppose $\{X_n : n \geq 1\}$ is a discrete time strictly stationary process with mixing constants $\{\alpha_n : n \geq 1\}$ such that for any positive integers k and n,

$$|P(A \cap B) - P(A)P(B)| \leq \alpha_n$$

for all sets A and B, where $A \in \sigma(X_1, \ldots, X_k)$, the σ-field generated by random variables $X_1, \ldots, X_k$, and $B \in \sigma(X_{k+n}, X_{k+n+1}, \ldots)$, the σ-field generated by random variables $X_{k+n}, X_{k+n+1}, \ldots$. If $\alpha_n \to 0$ as $n \to \infty$, then X_k and X_{k+n} are approximately independent for large n. In this case the sequence $\{X_n : n \geq 1\}$ is said to be strongly mixing (Rosenblatt [22], pp. 63–64).

Two discrete time strictly stationary processes $\{X_n : n \geq 1\}$ and $\{Y_n : n \geq 1\}$ are jointly strongly mixing with mixing constants $\{\alpha_n : n \geq 1\}$ if for arbitrary constants c_1 and c_2, let $Z_n = c_1 X_n + c_2 Y_n$, for each $n \geq 1$, then the process $\{Z_n : n \geq 1\}$ is strongly mixing with mixing constant $\{\alpha_n : n \geq 1\}$.

Let us define $S_n \equiv \sum_{i=1}^n X_i$ to be the partial sum, and $\overline{X}_n \equiv S_n/n$ to be the sample mean associated with the sequence. The subscript n will be omitted when there is no confusion. We now cite the following theorem.

24

Theorem 2.3 (Billingsley [5], p. 316). *Suppose that the sequence $\{X_n : n \geq 1\}$ is stationary and strongly mixing with $\alpha_n = O(n^{-5})$ and that $EX_n = 0$.*

(1) If $EX_n^4 < \infty$, then

$$n \operatorname{Var}[\overline{X}_n] \to \sigma^2 = EX_1^2 + 2 \sum_{k=1}^{\infty} EX_1 X_{k+1},$$

where the series converges absolutely.

(2) If $EX_n^{12} < \infty$ and $\sigma^2 > 0$, then $n^{1/2}\overline{X}_n/\sigma \Rightarrow N(0,1)$.

Corollary 2.4. *Suppose that the sequences $\{X_n : n \geq 1\}$ and $\{Y_n : n \geq 1\}$ are stationary and jointly strongly mixing with $\alpha_n = O(n^{-5})$ and that $EX_n = 0$, $EY_n = 0$, $EX_n^4 < \infty$, and $EY_n^4 < \infty$. Then*

$$n \operatorname{Cov}[\overline{X}_n, \overline{Y}_n] \to \sigma_{XY} = E[X_1 Y_1] + \sum_{k=1}^{\infty} EX_1 Y_{k+1} + \sum_{k=1}^{\infty} EY_1 X_{k+1},$$

where the series converges absolutely.

As noted in Billingsley [5], p. 316, in the proof of Theorem 2.3, the conditions $\alpha_n = O(n^{-5})$ and $EX_n^{12} < \infty$ are actually stronger than necessary; they are imposed to avoid technical complications in the proof. One can also see that the sufficient condition for $\operatorname{Var}[\overline{X}_n] = O(n^{-1})$ is weaker that that of the central limit theorem. Actually, as we will see in the next theorem, some general properties about the order of the cumulants of the sample mean can be obtained.

Theorem 2.5 (Titus [25], p. 16). *Consider a stationary sequence $\{X_n : n \geq 1\}$ such that: (1) for some positive integers j and k, $E|X_1|^{4(jk-1)} < \infty$, and (2) the*

sequence is mixing with $\alpha_n = O(n^{-2(j-1+\epsilon)})$ for some $\epsilon > 0$. Then for constants $a_0, \ldots, a_k$, $\kappa_j(\sum_{i=1}^k a_i S_n^i n^{1-i})$ is $O(n)$ as $n \to \infty$. Consequently, for constants $a_0, \ldots, a_k$, $\kappa_j(\sum_{i=1}^k a_i \overline{X}_n^i)$ is $O(n^{1-j})$ as $n \to \infty$.

Instead of the assumptions in Theorem 2.5 above, Titus [25] actually assumes that:

(1) For some positive integers j and k, each mixed moment of the form $E|X_{n_1} \cdots X_{n_l}|$ is bounded for all l, $1 \leq l \leq 4(jk-1)$, and (2) the sequence is mixing with $\alpha_n = O(n^{-2(j-1+\epsilon)})$ for some $\epsilon > 0$.

For a stationary process, his assumptions and ours are equivalent. It is easy to see that his assumption implies ours. The equivalency can be obtained by noting that, from Hölder's inequality, $E|X_{n_1} \cdots X_{n_l}| \leq (E|X_{n_1}|^l)^{1/l} \cdots (E|X_{n_l}|^l)^{1/l} = E|X_1|^l$, where the last equality follows from stationarity.

In both Theorems 2.3 and 2.5, the assumption that the sequence of mixing constants is decreasing fast enough is crucial. That assumption is true if, in addition, the strictly stationary process $\{X_n : n \geq 1\}$ is either independent, moving average of finite order, or m-dependent. Moreover, certain discrete time Markov chain has mixing constants that are decreasing exponentially with time; see Billingsley [5], p. 315, for details.

There are two more results that we will use latter on. The first one is a lemma from Titus [25].

Lemma 2.6 (Titus [25], p. 23). *Assume that $\{X_n : n \geq 1\}$ is a stationary sequence. Let $p = \sum_{i=1}^l p_i$, where each p_i is a positive integer. Suppose*

$E|X_1|^{4(p-1)} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-2(p+\epsilon)})$ for some $\epsilon > 0$. Then $\mathrm{cum}(S_n^{p_1}, \ldots, S_n^{p_l})$ is $O(n^{\min(\lfloor p/2 \rfloor, p-j+1)})$ as $n \to \infty$. Consequently, $\mathrm{cum}(\overline{X}_n^{p_1}, \ldots, \overline{X}_n^{p_l})$ is $O(n^{\min(-\lceil p/2 \rceil, 1-j)})$ as $n \to \infty$.

The second one is a corollary of Theorem 2.5. Suppose that $\kappa_i(S_n) = O(n)$ as $n \to \infty$ for each $i \le j$, then from properties (6) and (7) in Section 2.1, $\mu_i'(S_n) = O(n^i)$ and $\mu_i(S_n) = O(n^{\lfloor i/2 \rfloor})$ as $n \to \infty$ for each $i \le j$. From this observation, we have the following.

Corollary 2.7. Let $\{X_n : n \ge 1\}$ be a stationary sequence such that: (1) For some positive integers j and k, $E|X_1|^{4(jk-1)} < \infty$, (2) The sequence is mixing with $\alpha_n = O(n^{-2(j-1+\epsilon)})$ for some $\epsilon > 0$. Then for constants $a_0, \ldots, a_k$, $\mu_j'(\sum_{i=1}^k a_i S_n^i n^{1-i}) = O(n^j)$ and $\mu_j(\sum_{i=1}^k a_i S_n^i n^{1-i}) = O(n^{\lfloor j/2 \rfloor})$ as $n \to \infty$. Consequently, for constants $a_0, \ldots, a_k$, $\mu_j'(\sum_{i=1}^k a_i \overline{X}_n^i) = O(1)$ and $\mu_j(\sum_{i=1}^k a_i \overline{X}_n^i) = O(n^{-\lceil j/2 \rceil})$ as $n \to \infty$.

The cumulants of $\overline{X}_n$ can be calculated by the next lemma, which is also from Titus [25].

Lemma 2.8 (Titus [25], p. 29). Let $\{X_n : n \ge 1\}$ be a stationary process with mixing constants $\{\alpha_n : n \ge 1\}$.

(1) If $E|X_1|^4 < \infty$, and $\alpha_n = O(n^{-4-\epsilon})$ as $n \to \infty$ for some $\epsilon > 0$, then as $n \to \infty$. $\kappa_2(\overline{X}_n) = n^{-1} c_2 + n^{-2} d_2 + o(n^{-2})$, where

$$c_2 = \sum_{i=-\infty}^{\infty} \mathrm{cum}(X_0, X_i),$$

27

$$d_2 = -\sum_{i=1}^{\infty} 2i \ \mathrm{cum}(X_0, X_i).$$

(2) If $E|X_1|^8 < \infty$, and $\alpha_n = O(n^{-6-\epsilon})$ as $n \to \infty$ for some $\epsilon > 0$, then as $n \to \infty$,

$\kappa_3(\overline{X}_n) = n^{-2}c_3 + n^{-3}d_3 + o(n^{-3})$, where

$$c_3 = \sum_{i,j=-\infty}^{\infty} \mathrm{cum}(X_0, X_i, X_j),$$

$$d_3 = -\sum_{i=1}^{\infty} 3i \left[\mathrm{cum}(X_0, X_0, X_i) + \mathrm{cum}(X_0, X_i, X_i) \right]$$

$$- \sum_{i,j=1}^{\infty} 6(i+j) \mathrm{cum}(X_0, X_i, X_j).$$

(3) If $E|X_1|^{12} < \infty$, and $\alpha_n = O(n^{-8-\epsilon})$ as $n \to \infty$ for some $\epsilon > 0$, then as $n \to \infty$,

$\kappa_4(\overline{X}_n) = n^{-3}c_4 + n^{-4}d_4 + o(n^{-4})$, where

$$c_4 = \sum_{i,j,k=-\infty}^{\infty} \mathrm{cum}(X_0, X_i, X_j, X_k),$$

$$d_4 = -\sum_{i=1}^{\infty} 4i \left[\mathrm{cum}(X_0, X_0, X_0, X_i) + \mathrm{cum}(X_0, X_i, X_i, X_i) \right]$$

$$- \sum_{i=1}^{\infty} 6i \ \mathrm{cum}(X_0, X_0, X_i, X_i)$$

$$- \sum_{i,j=1}^{\infty} 12(i+j) \left[\mathrm{cum}(X_0, X_j, X_{i+j}, X_{i+j}) \right.$$

$$\left. + \mathrm{cum}(X_0, X_j, X_j, X_{i+j}) + \mathrm{cum}(X_0, X_0, X_j, X_{i+j}) \right]$$

$$- \sum_{i,j,k=1}^{\infty} 24(i+j+k) \mathrm{cum}(X_0, X_k, X_{j+k}, X_{i+j+k}).$$

28

Corollary 2.9. *Assume $\{X_n : n \geq 1\}$ is a stationary process with zero mean and S_n's are the partial sums:*

(1) If $\kappa_2(\overline{X}_n) = n^{-1} \cdot c_2 + n^{-2}d_2 + o(n^{-2})$, then $E(n^{1/2}\overline{X}_n)^2 = c_2 + n^{-1}d_2 + o(n^{-1})$.

(2) If $\kappa_i(\overline{X}_n) = n^{-i+1} \cdot c_i + n^{-i}d_i + o(n^{-i})$ for $2 \leq i \leq 3$, then $E(n^{1/2}\overline{X}_n)^3 = n^{-1/2}c_3 + n^{-3/2}d_3 + o(n^{-3/2})$.

(3) If $\kappa_i(\overline{X}_n) = n^{-i+1} \cdot c_i + n^{-i}d_i + o(n^{-i})$ for $2 \leq i \leq 4$, then $E(n^{1/2}\overline{X}_n)^4 = 3c_2^2 + n^{-1}[6c_2d_2 + c_4] + o(n^{-1})$.

(4) If $\kappa_i(\overline{X}_n) = n^{-i+1} \cdot c_i + n^{-i}d_i + o(n^{-i})$ for $2 \leq i \leq 5$, then $E(n^{1/2}\overline{X}_n)^5 = n^{-1/2}[10c_2c_3] + O(n^{-3/2})$.

(5) If $\kappa_i(\overline{X}_n) = n^{-i+1} \cdot c_i + n^{-i}d_i + o(n^{-i})$ for $2 \leq i \leq 6$, then $E(n^{1/2}\overline{X}_n)^6 = 15c_2^3 + O(n^{-1})$.

2.6. Consistency of Sample Moments

Let $\{X_i : i \geq 1\}$ be a discrete time strictly stationary process. Assume that $\{\mu'_r : r \geq 1\}$ and $\{\mu_r : r \geq 1\}$ are the moments and central moments of X_1. The $\{\mu'_r : r \geq 1\}$ and $\{\mu_r : r \geq 1\}$ represent important parameters about the distribution of X_1. Natural estimators of these parameters are given by the corresponding sample moments of $\{X_i : i \geq 1\}$. Thus μ'_r may be estimated by

$$m'_{n,r} = (1/n) \sum_{i=1}^{n} X_i^r, \tag{2.59}$$

and μ_r may be estimated by

$$m_{n,r} = (1/n) \sum_{i=1}^{n} (X_i - \overline{X})^r, \tag{2.60}$$

29

for each $r \geq 1$. For convenience, we also denote $\mu = \mu'_1 = EX_1$ and $\overline{X}_n = m'_{n,1}$.
In this section, we follow the development of Serfling [24], pp. 67-71, to show the mean square consistency of these estimators. We begin with the following.

Proposition 2.10. *Assume the sequence $\{X_i : i \geq 1\}$ is stationary and strongly mixing with $\alpha_i = O(i^{-5})$ and, for an integer r, $EX_1^{4r} < \infty$. Then*

(1) $Em'_{n,r} = \mu'_r$;

(2) $\text{Var}\{m'_{n,r}\} = O(n^{-1})$;

(3) $\text{MSE}\{m'_{n,r}\} \equiv E\{m'_{n,r} - \mu'_r\}^2 = O(n^{-1})$;

(4) $m'_{n,r} \to \mu'_r$ in L^2 and in distribution, as $n \to \infty$.

Proof: (1) Since $EX_i^r = \mu'_r$ for each i, the result follows.

(2) Notice that the process $\{X_i^r : i \geq 1\}$ is stationary and strongly mixing with $\alpha_i = O(i^{-5})$; this result then follows from Theorem 2.3.

(3) This result follows from (1) and (2).

(4) This result follows from (3). $\qquad\qquad\qquad\qquad$ □

Preliminary to stating properties of the estimates $m_{n,r}$'s, it is advantageous to consider the closely related random variables

$$m^*_{n,r} = (1/n) \sum_{i=1}^{n} (X_i - \mu)^r, \qquad\qquad (2.61)$$

for each $r \geq 1$. Properties of $m_{n,r}$'s will be deduced from those of the $m^*_{n,r}$'s. The same arguments employed in Proposition 2.10 immediately yield the following.

Proposition 2.11. *Assume the sequence $\{X_i : i \geq 1\}$ is stationary and strongly mixing with $\alpha_i = O(i^{-5})$ and, for an integer r, $EX_1^{4r} < \infty$. Then*

30

(1) $Em_{n,r}^* = \mu_r$;

(2) $\text{Var}\{m_{n,r}^*\} = O(n^{-1})$;

(3) $\text{MSE}\{m_{n,r}^*\} \equiv E\{m_{n,r}^* - \mu_r\}^2 = O(n^{-1})$;

(4) $m_{n,r}^* \to \mu_r$ in L^2 and in distribution, as $n \to \infty$.

We record the following equation from Serfling [24], p. 69.

$$m_{n,r} = (1/n) \sum_{i=1}^{n} (X_i - \overline{X})^r$$

$$= (1/n) \sum_{i=1}^{n} \sum_{j=0}^{r} \binom{r}{j} (X_i - \mu)^j (\mu - \overline{X})^{r-j}$$

$$= \sum_{j=0}^{r} \binom{r}{j} (-1)^{r-j} m_{n,j}^* (m_{n,1}^*)^{r-j}, \qquad (2.62)$$

where we define $m_{n,0}^* = 1$. Although analogous in form to $m_{n,r}^*$'s, the random variables $m_{n,r}$'s are much harder to analyze. Therefore, instead of dealing with $m_{n,r}$'s directly, we exploit the relationship between $m_{n,r}$'s and $m_{n,r}^*$'s. We have the following proposition.

Proposition 2.12. *Assume the sequence* $\{X_i : i \geq 1\}$ *is stationary and strongly mixing with* $\alpha_i = O(i^{-5})$ *and, for an integer* r, $EX_1^{4r} < \infty$. *If, in addition,* $E(m_1^*)^j = O(n^{-\lceil j/2 \rceil})$, *for each* j, $0 \leq j \leq r$, *then*

(1) $Em_{n,r} = \mu_r + O(n^{-1})$;

(2) $\text{Var}\{m_{n,r}\} = O(n^{-1})$;

(3) $\text{MSE}\{m_{n,r}\} \equiv E\{m_{n,r} - \mu_r\}^2 = O(n^{-1})$;

(4) $m_{n,r} \to \mu_r$ in L^2 and in distribution, as $n \to \infty$.

31

Proof: (1) Utilize equation (2.62) to write

$$Em_{n,r} - \mu_r = \sum_{j=1}^{r} \binom{r}{j} (-1)^j E\{m_{n,r-j}^* (m_{n,1}^*)^j\}.$$

Now, observe that

$$E\{m_{n,r-1}^* m_{n,1}^*\} = (1/n^2) E\left\{ \sum_{i=1}^{n} (X_i - \mu)^{r-1} \sum_{j=1}^{n} (X_j - \mu) \right\}$$

$$= (1/n) \sum_{i=-\infty}^{\infty} E\{(X_0 - \mu)^{r-1} (X_i - \mu)\} + O(n^{-2})$$

$$= O(n^{-1}),$$

where the second equality is obtained from the fact that $\{(X_i - \mu) : i \geq 1\}$ is zero mean, $\{(X_i - \mu)^{r-1} : i \geq 1\}$ and $\{(X_i - \mu) : i \geq 1\}$ are jointly strongly mixing, the fourth central moments of both $(X_1 - \mu)^{r-1}$ and $(X_1 - \mu)$ are both finite, and then by an application of Corollary 2.4. For each $j \geq 2$, use Hölder's inequality

$$|E\{m_{n,r-j}^* (m_{n,1}^*)^j\}| \leq \left[E|m_{n,r-j}^*|^{r/(r-j)} \right]^{(r-j)/r} \left[E|m_{n,1}^*|^r \right]^{j/r}.$$

By application of Minkowski's inequality

$$\left[E|m_{n,r-j}^*|^{r/(r-j)} \right]^{(r-j)/r} = \left[E\left\{ |(1/n) \sum_{i=1}^{n} (X_i - \mu)^{r-j}|^{r/(r-j)} \right\} \right]^{(r-j)/r}$$

$$\leq (1/n) \sum_{i=1}^{n} \left[E\{|X_i - \mu|^r\} \right]^{(r-j)/r}$$

$$= O(1),$$

and the assumption that $E(m_{n,1}^*)^r = O(n^{-\lceil r/2 \rceil})$, we obtain

$$E\{m_{n,r-j}^* (m_{n,1}^*)^j\} = O(1) \left[O(n^{-\lceil r/2 \rceil}) \right]^{j/r}$$

32

$$= O(1)O(n^{-j/2})$$

$$= O(n^{-1}).$$

The results then follow.

(2) Writing $\text{Var}\{m_{n,r}\} = E\{m_{n,r}^2\} - \{Em_{n,r}\}^2$, we seek to compute $E\{m_{n,r}^2\}$ and combine the result in (1). To do this, we need to compute quantities of the form $E\{m_{n,r-j_1}^*(m_{n,1}^*)^{j_1} m_{n,r-j_2}^*(m_{n,1}^*)^{j_2}\} = E\{m_{n,r-j_1}^* m_{n,r-j_2}^*(m_{n,1}^*)^{j_1+j_2}\}$, for each j_1 and j_2, $0 \le j_1, j_2 \le r$. For $j_1 = j_2 = 0$, we have $E(m_{n,r}^*)^2 = \kappa_2\{m_{n,r}^*\} + \{Em_{n,r}\}^2$. But

$$\kappa_2\{m_{n,r}^*\}$$

$$= (1/n^2)E\left\{\left[\sum_{i=1}^{n}[(X_i - \mu)^r - \mu_r]\right]^2\right\}$$

$$= (1/n)\sum_{i=-\infty}^{\infty} E\left\{\left[(X_0 - \mu)^r - \mu_r\right]\left[(X_i - \mu)^r - \mu^r\right]\right\} + O(n^{-2})$$

$$= O(n^{-1}),$$

where the equalities are justified by the moment and mixing conditions as well as an application of Corollary 2.4. For $(j_1, j_2) = (0, 1)$ or $(1, 0)$, we have

$$E\{m_{n,r-1}^* m_{n,r}^* m_{n,1}^*\} = \text{cum}(m_{n,r-1}^*, m_{n,r}^*, m_{n,1}^*)$$

$$+ \text{cum}(m_{n,r-1}^*)\text{cum}(m_{n,r}^*, m_{n,1}^*)$$

$$+ \text{cum}(m_{n,r}^*)\text{cum}(m_{n,r-1}^*, m_{n,1}^*)$$

$$+ \text{cum}(m_{n,1}^*)\text{cum}(m_{n,r-1}^*, m_{n,r}^*)$$

$$+ \text{cum}(m_{n,r-1}^*)\text{cum}(m_{n,r}^*)\text{cum}(m_{n,1}^*).$$

The first term is at most $O(n^{-2})$ by Theorem 2.5. Each of the next two terms is at most $O(n^{-1})$, which can be proved in a similar fashion as in (1). The last two terms on the right hand side are both zero since $Em_{n,1}^* = 0$. Thus $E\{m_{n,r-1}^* m_{n,r}^* m_{n,1}^*\} =$

$O(n^{-1})$. Finally, for $j_1 + j_2 \geq 2$, by a similar application of the Hölder's inequality and the Minkowski's inequality, we have

$$E\{m^*_{n,r-j_1} m^*_{n,r-j_2} (m^*_{n,1})^{j_1+j_2}\} = O(n^{-1}).$$

The desired result now follows.

(3) This result follows from (1) and (2).

(4) This result follows from (3). □

Chapter 3

Johnson-Glynn Pivots for Stationary Processes

In order to obtain a more accurate confidence interval it is necessary that we first obtain the formal Edgeworth and Cornish-Fisher expansions of the sample statistic of interest. However, to obtain these expansions we have to calculate and estimate various moments and cumulants of the sample statistic. In this chapter, we will use this approach to obtain the first and second order Johnson-Glynn pivots and the associated confidence intervals for the batch means method.

The organization of this chapter is as follows. In Section 3.1, we will review the traditional batch means method. Section 3.2 shows the calculation of some mixed cumulants related to the batch means method. The cumulants of the sample t-statistic and the first and second order Johnson-Glynn pivots for the batch means method will be presented in Section 3.3. Section 3.4 derives two new confidence intervals from the Johnson-Glynn pivots and shows that the increase of the length of the confidence intervals are asymptotically negligible. Section 3.5 compares the computational efficiency of the traditional batch means method with that of the first and second order Johnson-Glynn pivots.

3.1. Batch Means Method

Suppose $\{X_i : i \geq 1\}$ is a discrete time strictly stationary process. We are interested in obtaining a point estimate and a confidence interval for the stationary mean EX, where X has the same distribution as X_1. In this section, we will review

35

the traditional batch means method. Without loss of generality, we assume that $\{X_i : i \geq 1\}$ has zero-mean.

To show that a confidence interval for EX can be constructed, we follow the development of Brillinger [6]. However, our assumptions about the stationary processes are slightly different. Specifically, Brillinger ([6], p. 419) assumes that:

> The discrete time process $\{X_i : i \geq 1\}$ is stationary, continuous in mean, and for each $k, i \geq 1$, and $j_1, \ldots, j_k \geq 0$, its cumulant of order $k + 1$ exists and satisfies $|\text{cum}\{X_i, X_{i+j_1}, \ldots, X_{i+j_k}\}| < L_k(1 + j_1^2)^{-1} \cdots (1 + j_k^2)^{-1}$, for some finite positive constant L_k.

In our discussion, instead of the assumption that the cumulant functions decrease reasonably fast as the arguments become far apart, we assume that the stationary processes satisfy certain moment conditions and are strongly mixing with the mixing constants α_i's decrease as $i \rightarrow \infty$. Notice that both assumptions require that sufficient number of the moments of the random variable X are finite and also require that the values of the stationary process at a distance from each other are only weakly statistically dependent. Both assumptions are true if, in addition to their respective moment conditions, the strictly stationary process $\{X_i : i \geq 1\}$ is either independent, moving average of finite order, or m-dependent.

Notice that in the following discussion, only discrete time strictly stationary processes are considered; however, the results for continuous time processes can be obtained in a similar fashion (see Brillinger [6]).

Let

$$\overline{X}_m \equiv (1/m) \sum_{i=1}^{m} X_i \tag{3.1}$$

be the sample mean. From Theorem 2.3, we know that if the mixing constants α_i are $O(i^{-5})$ as $i \to \infty$, $EX_i^{12} < \infty$, and $\sigma^2 = EX_1^2 + 2\sum_{i=1}^{\infty} EX_1 X_{i+1} > 0$, then $\overline{X}_m/(\sigma/m^{1/2})$ converges weakly to $N(0,1)$. Thus if we can estimate σ then we can use $\overline{X}_m$ as a point estimate and construct a confidence interval for EX.

Suppose the time interval $[1, m]$ is split into n intervals of length $b = \lfloor m/n \rfloor$ each and Y_i denotes the mean for the ith interval, i.e.

$$Y_i = \frac{1}{b} \sum_{j=(i-1)b+1}^{ib} X_j, \qquad (3.2)$$

for $i = 1, \ldots, n$. For the Batch Means Method, n is the number of batches, b is the batch size, and the Y_i's are batch means. The basic idea for batching is as follows. We have a stationary process whose values at a distance are only weakly statistically dependent. By batching, we transform the original data sequence to another one and, hopefully, the dependency of that sequence will decay faster. As a matter of fact, we can show that the batch means are asymptotically independent, in the sense that $|P(A \cap B) - P(A)P(B)| \to 0$ for all sets A and B, where $A \in \sigma(Y_1, \ldots, Y_k)$, and $B \in \sigma(Y_{k+n}, \ldots)$, as $n \to \infty$. We state the following proposition, whose proof will be given at the Appendix.

Proposition 3.1. *Suppose $\{X_i : i \geq 1\}$ is a discrete time, strictly stationary stochastic process. For a fixed positive integer b and a real function $g: \mathcal{R}^b \to \mathcal{R}$, where $\mathcal{R}$ is the set of real numbers, let $Y_i = g(X_{(i-1)b+1}, X_{(i-1)b+2}, \ldots, X_{ib})$, for each $i \geq 1$. Then*

(1) The process $\{Y_i : i \geq 1\}$ is strictly stationary.

(2) Suppose, in addition to being stationary, $\{X_i : i \geq 1\}$ is strongly mixing with

mixing constants $\{\alpha_i : i \geq 1\}$. *Without loss of generality, we can assume that the sequence of mixing constants $\{\alpha_i : i \geq 1\}$ is a nonnegative and nonincreasing sequence of i. Then $\{Y_i : i \geq 1\}$ is strongly mixing with mixing constants $\{\beta_i : i \geq 1\}$, where for each $i \geq 1$, $\beta_i \equiv \alpha_{(i-1)b+1}$.*

(3) For an increasing function $h: \mathcal{R}^+ \to \mathcal{R}^+$, where $\mathcal{R}^+$ is the set of nonnegative real numbers, $\sum_{i=1}^{\infty} h(\alpha_i) < \infty$ if and only if $\sum_{i=1}^{\infty} h(\beta_i) < \infty$.

(4) Suppose, in particular, $Y_i \equiv (1/b) \sum_{j=1}^{b} X_{(i-1)b+j}$ for each $i \geq 1$. Then $E|X_1|^j < \infty$ implies that $E|Y_1|^j < \infty$.

Basically, Proposition 3.1 states that the stationarity, mixing, and moment conditions of a stationary process, as well as the finite summability of the mixing constants are preserved after batching. We then have the following result.

Theorem 3.2 (Brillinger [6], p. 420). *Suppose $\{X_i : i \geq 1\}$ is stationary and strongly mixing with $\alpha_i = O(i^{-5})$ and that $EX_i = 0$ and $EX_i^{12} < \infty$. Then for fixed n, the Y_i's are asymptotically independent and individually asymptotically normal with mean EX and variance*

$$(1/b)\sigma^2 = (1/b)\{EX_1^2 + 2\sum_{k=1}^{\infty} EX_1 X_{k+1}\},$$

as $m \to \infty$.

Since the Y_i's are asymptotically independent and and individually asymptotically normal, the next natural question is whether we can use the usual approach in analyzing i.i.d. normal random samples to generate a confidence interval for $\overline{Y}$.

which is $\overline{X}_m$. Define

$$V^*_{n,b} \equiv \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \overline{X}_m)^2. \tag{3.3}$$

as the sample variance for $\{Y_i\}$ and let

$$t^*_{n,b} \equiv \frac{\overline{X}_m - EX}{(V^*_n/n)^{1/2}}. \tag{3.4}$$

We have the following.

Corollary 3.3 (Brillinger [6], p. 421). *Suppose $\{X_i : i \geq 1\}$ is stationary and strongly mixing with $\alpha_i = O(i^{-5})$ and that $EX_i = 0$ and $EX_i^{12} < \infty$. Then for fixed n, the random variable $t^*_{n,b}$ converges to a Student-t random variable with $n-1$ degrees of freedom in the limit as $m \to \infty$.*

Since a Student-t random variable with $n-1$ degrees of freedom converges to $N(0,1)$ as $n \to \infty$, we can use the quantile points of the Student-t random variable or the standard normal random variable to construct a confidence interval for EX.

3.2. Cumulants in Batch Means Method

The basic idea of the batch means method is to choose the batch size large enough so that, hopefully, the correlation between the batch means becomes small and each batch means is reasonably close to a normal. Then the confidence interval methods for the i.i.d. normal random variables can be applied (see Corollary 3.3). Thus the properties such as joint cumulants and moments or the mixing constants of the batch means are important in the batch means method. In this section, various cumulants related to the batch means will be given.

Recall that for each i, $1 \leq i \leq n$, $Y_i = (1/b) \sum_{j=1}^{b} X_{(i-1)b+j}$ is the ith batch mean. We also define

$$V_{n,b} \equiv \frac{1}{n} \sum_{i=1}^{n} (Y_i - \overline{X}_m)^2, \qquad (3.5)$$

$$\sigma_m^2 \equiv m \cdot \text{Var}(\overline{X}_m), \qquad (3.6)$$

and

$$\Delta_{n,b} \equiv \frac{V_{n,b}}{(\sigma_m^2/b)} - 1. \qquad (3.7)$$

We will compute the order of various mixed cumulants of $\overline{X}_m$, $\overline{X}_m^2$, $(1/n) \sum_{i=1}^{n} Y_i^2$, and $\Delta_{n,b}$. In the following derivation, we will assume that the process $\{X_i : 1 \leq i \leq m\}$ satisfies sufficient mixing and moment conditions so that there will not be any problem regarding convergence of infinite sums. These conditions can be obtained by applying Theorem 2.5. Below we list the asymptotic order of the cumulants, while relegating the derivation to the Appendix. Each of the results below requires certain order conditions on mixing constants and moments. As these are quite complicated, they will only be given in the Appendix.

Cumulants of $\overline{X}_m$

$\kappa_2(\overline{X}_m) = O(m^{-1})$

$\kappa_3(\overline{X}_m) = O(m^{-2})$

$\kappa_4(\overline{X}_m) = O(m^{-3})$

$\kappa_5(\overline{X}_m) = O(m^{-4})$

Cumulants of $\overline{X}_m^2$

$\kappa_1(\overline{X}_m^2) = O(m^{-1})$

$$\kappa_2(\overline{X}_m^2) = O(m^{-2})$$

$$\kappa_3(\overline{X}_m^2) = O(m^{-3})$$

$$\kappa_4(\overline{X}_m^2) = O(m^{-4})$$

Mixed Cumulants of $\overline{X}_m$ and $\overline{X}_m^2$

$$\kappa_{1,1}(\overline{X}_m, \overline{X}_m^2) = O(m^{-2})$$

$$\kappa_{2,1}(\overline{X}_m, \overline{X}_m^2) = O(m^{-2})$$

$$\kappa_{1,2}(\overline{X}_m, \overline{X}_m^2) = O(m^{-3})$$

$$\kappa_{3,1}(\overline{X}_m, \overline{X}_m^2) = O(m^{-3})$$

$$\kappa_{2,2}(\overline{X}_m, \overline{X}_m^2) = O(m^{-3})$$

$$\kappa_{1,3}(\overline{X}_m, \overline{X}_m^2) = O(m^{-4})$$

$$\kappa_{4,1}(\overline{X}_m, \overline{X}_m^2) = O(m^{-4})$$

$$\kappa_{3,2}(\overline{X}_m, \overline{X}_m^2) = O(m^{-4})$$

$$\kappa_{2,3}(\overline{X}_m, \overline{X}_m^2) = O(m^{-4})$$

$$\kappa_{5,1}(\overline{X}_m, \overline{X}_m^2) = O(m^{-5})$$

$$\kappa_{4,2}(\overline{X}_m, \overline{X}_m^2) = O(m^{-5})$$

$$\kappa_{3,3}(\overline{X}_m, \overline{X}_m^2) = O(m^{-5})$$

$$\kappa_{6,1}(\overline{X}_m, \overline{X}_m^2) = O(m^{-6})$$

$$\kappa_{5,2}(\overline{X}_m, \overline{X}_m^2) = O(m^{-6})$$

$$\kappa_{7,1}(\overline{X}_m, \overline{X}_m^2) = O(m^{-7})$$

Cumulants of $(1/n)\sum_{i=1}^n Y_i^2$

$$\kappa_1((1/n)\sum_{i=1}^n Y_i^2) = O(b^{-1})$$

$$\kappa_2((1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-1})O(b^{-2})$$

$$\kappa_3((1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-2})O(b^{-3})$$

$$\kappa_4((1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-3})O(b^{-4})$$

Mixed cumulants of $\overline{X}_m$ and $(1/n)\sum_{i=1}^{n} Y_i^2$

$$\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-1})O(b^{-2})$$

$$\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-2})O(b^{-2})$$

$$\kappa_{1,2}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-2})O(b^{-3})$$

$$\kappa_{3,1}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-3})O(b^{-3})$$

$$\kappa_{2,2}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-3})O(b^{-3})$$

$$\kappa_{1,3}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-3})O(b^{-4})$$

$$\kappa_{4,1}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-4})O(b^{-4})$$

$$\kappa_{3,2}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-4})O(b^{-4})$$

$$\kappa_{2,3}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-4})O(b^{-4})$$

$$\kappa_{5,1}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-5})O(b^{-5})$$

$$\kappa_{4,2}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-5})O(b^{-5})$$

$$\kappa_{3,3}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-5})O(b^{-5})$$

$$\kappa_{6,1}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-6})O(b^{-6})$$

$$\kappa_{5,2}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-6})O(b^{-6})$$

$$\kappa_{7,1}(\overline{X}_m, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-7})O(b^{-7})$$

Mixed cumulants of $\overline{X}_m^2$ and $(1/n)\sum_{i=1}^{n} Y_i^2$

$$\kappa_{1,1}(\overline{X}_m^2, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-2})O(b^{-2})$$

$$\kappa_{2,1}(\overline{X}_m^2, (1/n)\sum_{i=1}^{n} Y_i^2) = O(n^{-3})O(b^{-3})$$

42

$$\kappa_{1,2}(\overline{X}_m^2, (1/n) \sum_{i=1}^n Y_i^2) = O(n^{-3})O(b^{-3}) + O(n^{-2})O(b^{-4})$$

Mixed cumulants of $\overline{X}_m$, $\overline{X}_m^2$, and $(1/n) \sum_{i=1}^n Y_i^2$

$$\kappa_{1,1,1}(\overline{X}_m, \overline{X}_m^2, (1/n) \sum_{i=1}^n Y_i^2) = O(n^{-2})O(b^{-3})$$

$$\kappa_{2,1,1}(\overline{X}_m, \overline{X}_m^2, (1/n) \sum_{i=1}^n Y_i^2) = O(n^{-3})O(b^{-3})$$

$$\kappa_{1,2,1}(\overline{X}_m, \overline{X}_m^2, (1/n) \sum_{i=1}^n Y_i^2) = O(n^{-3})O(b^{-4})$$

$$\kappa_{1,1,2}(\overline{X}_m, \overline{X}_m^2, (1/n) \sum_{i=1}^n Y_i^2) = O(n^{-3})O(b^{-4})$$

$$\kappa_{3,1,1}(\overline{X}_m, \overline{X}_m^2, (1/n) \sum_{i=1}^n Y_i^2) = O(n^{-4})O(b^{-4})$$

$$\kappa_{2,2,1}(\overline{X}_m, \overline{X}_m^2, (1/n) \sum_{i=1}^n Y_i^2) = O(n^{-4})O(b^{-4})$$

$$\kappa_{2,1,2}(\overline{X}_m, \overline{X}_m^2, (1/n) \sum_{i=1}^n Y_i^2) = O(n^{-4})O(b^{-4})$$

$$\kappa_{4,1,1}(\overline{X}_m, \overline{X}_m^2, (1/n) \sum_{i=1}^n Y_i^2) = O(n^{-5})O(b^{-5})$$

$$\kappa_{3,2,1}(\overline{X}_m, \overline{X}_m^2, (1/n) \sum_{i=1}^n Y_i^2) = O(n^{-5})O(b^{-5})$$

$$\kappa_{3,1,2}(\overline{X}_m, \overline{X}_m^2, (1/n) \sum_{i=1}^n Y_i^2) = O(n^{-5})O(b^{-5})$$

Cumulants of $\Delta_{n,b}$

$$\kappa_1(\Delta_{n,b}) = O(b^{-1}) + O(n^{-1})$$

$$\kappa_2(\Delta_{n,b}) = O(n^{-1})$$

$$\kappa_3(\Delta_{n,b}) = O(n^{-2})$$

Mixed cumulants of $\Delta_{n,b}$ and $\overline{X}_m$

$$\kappa_{1,1}(\Delta_{n,b}, \overline{X}_m) = O(n^{-1})O(b^{-1})$$

$$\kappa_{2,1}(\Delta_{n,b}, \overline{X}_m) = O(n^{-2})O(b^{-1})$$

$$\kappa_{1,2}(\Delta_{n,b}, \overline{X}_m) = O(n^{-2})O(b^{-1})$$

$$\kappa_{3,1}(\Delta_{n,b}, \overline{X}_m) = O(n^{-3})O(b^{-1})$$

$$\kappa_{2,2}(\Delta_{n,b}, \overline{X}_m) = O(n^{-3})O(b^{-1})$$

$$\kappa_{1,3}(\Delta_{n,b}, \overline{X}_m) = O(n^{-3})O(b^{-2})$$

$$\kappa_{3,2}(\Delta_{n,b}, \overline{X}_m) = O(n^{-4})O(b^{-1})$$

$$\kappa_{2,3}(\Delta_{n,b}, \overline{X}_m) = O(n^{-4})O(b^{-2})$$

$$\kappa_{1,4}(\Delta_{n,b}, \overline{X}_m) = O(n^{-4})O(b^{-3})$$

$$\kappa_{3,3}(\Delta_{n,b}, \overline{X}_m) = O(n^{-5})O(b^{-2})$$

$$\kappa_{2,4}(\Delta_{n,b}, \overline{X}_m) = O(n^{-5})O(b^{-3})$$

3.3. Johnson-Glynn Pivots

Suppose a discrete time strictly stationary process $\{X_i : i \geq 1\}$ is strongly mixing with mixing constants $\{\alpha_i : i \geq 1\}$. We are interested in applying the ideas of the Johnson-Glynn pivots to obtain a confidence interval for the stationary mean EX, where X has the same distribution as X_i. Again, we assume that $\{X_i : i \geq 1\}$ is zero-mean.

Our starting point is the batch means method, which we intend to modify to obtain new confidence interval methods. The basic idea about batch means method has been discussed in Section 3.1. Recall that, for each i, $1 \leq i \leq n$, $Y_i \equiv (1/b)\sum_{j=1}^{b} X_{(i-1)b+j}$, $V_{n,b} \equiv (1/n)\sum_{i=1}^{n}(Y_i - \overline{X}_m)^2$, $\sigma_m^2 \equiv m \cdot \mathrm{Var}(\overline{X}_m)$, and $\Delta_{n,b} \equiv V_{n,b}/(\sigma_m^2/b) - 1$; see equations (3.2), (3.5), (3.6), and (3.7). We also define

$$t_{n,b} \equiv \frac{\overline{X}_m - EX}{(V_{n,b}/n)^{1/2}}. \tag{3.8}$$

In this section, we will demonstrate how to calculate the first four cumulants of

44

$t_{n,b}$.

To handle various cumulants of $t_{n,b}$, note that, from equation (3.7), $V_{n,b} = (\sigma_m^2/b)\{1 + \Delta_{n,b}\}$. From Lemma A.12 in the Appendix, $\Delta_{n,b}$ converges to 0 in L^2, as $n, b \to \infty$. This implies that $\Delta_{n,b}$ is small for large n and b. Simple algebra and a Taylor expansion then yields

$$\frac{1}{V_{n,b}} = \frac{1}{\sigma_m^2/b}\left\{1 - \Delta_{n,b} + \Delta_{n,b}^2 - \Delta_{n,b}^3 + \cdots\right\}, \qquad (3.9)$$

so that, from equation (3.8),

$$t_{n,b} = n^{1/2}(\overline{X}_m - EX)\left(\frac{1}{V_{n,b}}\right)^{1/2}$$

$$= \frac{m^{1/2}}{\sigma_m}(\overline{X}_m - EX)\left\{1 - \frac{1}{2}\Delta_{n,b} + \frac{3}{8}\Delta_{n,b}^2 - \frac{5}{16}\Delta_{n,b}^3 + \cdots\right\}. \qquad (3.10)$$

It is clear that if we can calculate the various mixed cumulants of $(\overline{X}_m - EX)$, $(\overline{X}_m - EX)\Delta_{n,b}$, $(\overline{X}_m - EX)\Delta_{n,b}^2$, ..., etc., then we have obtained various cumulants of $t_{n,b}$.

The whole procedure is very technical and the exact results are very involved. Basically, the idea is as follows. Notice that we assume that $\{X_i : 1 \le i \le m\}$ is zero-mean. We need to calculate various mixed cumulants of $\overline{X}_m$, $\overline{X}_m\Delta_{n,b}$, $\overline{X}_m\Delta_{n,b}^2$, $\overline{X}_m\Delta_{n,b}^3$, ..., etc. But those cumulants can be obtained after calculating various mixed cumulants of $\overline{X}_m$ and $\Delta_{n,b}$. Next, observe that from equation (3.7). $\Delta_{n,b} = [V_{n,b}/(\sigma_m^2/b)] - 1$, but from equation (3.5), $V_{n,b} \equiv (1/n)\sum_{i=1}^n (Y_i - \overline{X}_m)^2 = (1/n)\sum_{i=1}^n Y_i^2 - \overline{X}_m^2$, so that we have to calculate mixed cumulants of $\overline{X}_m$, $\overline{X}_m^2$, and $(1/n)\sum_{i=1}^n Y_i^2$, which in turn can be solved by first calculating mixed cumulants

of $\overline{X}_m$ and $(1/n)\sum_{i=1}^n Y_i^2$. Notice that the order of each of these cumulants and mixed cumulants mentioned above have been discussed in Section 3.2. Applying these ideas, we derive the following.

Proposition 3.4. *Assume that $\{X_i : i \geq 1\}$ is a discrete time, strictly stationary stochastic process with zero mean. Then*

(1) If $E|X_1|^8 < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-7})$, then $\kappa_1(t_{n,b}) =$

$-(1/2)(\sqrt{m}b/\sigma_m^3)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + o(n^{-1/2})o(b^{-1/2})$.

(2) If $E|X_1|^{12} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-9})$, then $\kappa_2(t_{n,b}) =$

$1 + (b/\sigma_m^2)^2\kappa_2((1/n)\sum_{i=1}^n Y_i^2) - (mb/\sigma_m^2)\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)$

$- E[(1/n)\sum_{i=1}^n Y_i^2/(\sigma_m^2/b) - 1] + 3/n + o(n^{-1}) + o(b^{-1})$.

(3) If $E|X_1|^{16} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-11})$, then $\kappa_3(t_{n,b}) =$

$(\sqrt{m}/\sigma_m)^3\{\kappa_3(\overline{X}_m) - 3(\sigma_m^2/m)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2/(\sigma_m^2/b))\} + o(n^{-1/2})o(b^{-1/2})$.

(4) If $E|X_1|^{20} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-13})$, then $\kappa_4(t_{n,b}) =$

$-6(mb/\sigma_m^4)\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 3(b/\sigma_m^2)^2\kappa_2((1/n)\sum_{i=1}^n Y_i^2) + 12/n + o(n^{-1})$.

It can also be shown that we have $\kappa_1(t_{n,b}) = O(n^{-1/2})O(b^{-1/2})$, $\kappa_2(t_{n,b}) = 1 + O(n^{-1}) + O(b^{-1})$, $\kappa_3(t_{n,b}) = O(n^{-1/2})O(b^{-1/2})$, and $\kappa_4(t_{n,b}) = O(n^{-1})$. We also note that the $O(b^{-1})$ term in $\kappa_2(t_{n,b})$ would disappear should $\{Y_i : 1 \leq i \leq m\}$ be independent.

The first four cumulants of a standard normal random variable are 0, 1, 0, and 0. If we want to use a normal approximation to generate a confidence interval, we would like to have the cumulants of $t_{n,b}$ to be as close to those of the standard normal random variable as possible.

This idea follows from Edgeworth expansion in the following sense. In the

46

Edgeworth expansion for the distribution function of $t_{n,b}$, the error terms have coefficients in terms of the cumulants of $t_{n,b}$, namely, $\kappa_1(t_{n,b})$, $\kappa_2(t_{n,b}) - 1$, $\kappa_3(t_{n,b})$, and $\kappa_4(t_{n,b})$; see equation (2.31). Thus if $\kappa_1(t_{n,b})$, $\kappa_2(t_{n,b}) - 1$, $\kappa_3(t_{n,b})$, and $\kappa_4(t_{n,b})$ are close to zero, the effects of the error terms will be small. Notice that the first four cumulants of a standard normal random variable are 0, 1, 0, and 0. We remark parenthetically that requiring $\kappa_1(t_{n,b}) \to 0$, $\kappa_2(t_{n,b}) - 1 \to 0$, $\kappa_3(t_{n,b}) \to 0$. and $\kappa_4(t_{n,b}) \to 0$ is the same as requiring that the first four cumulants of $t_{n,b}$ are close to those of a standard normal random variable.

The above argument leads to two naive, but natural, approaches for the choice of the batch size: (1) Minimizing $\max_{1 \leq i \leq 4} |\kappa_i(t_{n,b}) - \kappa_i(\xi)|$. (2) Minimizing $\sum_{i=1}^{4} [\kappa_i(t_{n,b}) - \kappa_i(\xi)]^2$. From Proposition 3.4, $\kappa_1(t_{n,b}) - \kappa_1(\xi) = O(n^{-1/2})O(b^{-1/2})$, $\kappa_2(t_{n,b}) - \kappa_2(\xi) = O(n^{-1}) + O(b^{-1})$, $\kappa_3(t_{n,b}) - \kappa_3(\xi) = O(b^{-1})$, and $\kappa_4(t_{n,b}) - \kappa_4(\xi) = O(n^{-1})$. In both cases we can see that the optimal b and n should be chosen such that b and n are of the same order, namely, both $b \sim O(m^{1/2})$ and $n \sim O(m^{1/2})$ as $m \to \infty$. This relationship of the batch size and the number of batches gives us a $t_{n,b}$ which is closer to a standard normal random variable in the sense that the differences between their respective first four cumulants are smaller. Hence we feel this is a preferred choice of batch size for the traditional batch means method.

After calculating various cumulants, the formal Edgeworth expansion and Cornish-Fisher expansion can be obtained. Notice that for fixed m, the choice of b and n will affect both formal expansions. For $n \sim O(m^{1/2})$ and $b \sim O(m^{1/2})$ the formal Edgeworth expansion is

$$F_{t_{n,b}}(x) = \Phi(x) + \left(\frac{\kappa_3(t_{n,b})}{6} - \kappa_1(t_{n,b}) \right)\phi(x) + \left(\frac{\kappa_4(t_{n,b})}{8} - \frac{\kappa_2(t_{n,b}) - 1}{2} \right)x\phi(x)$$

$$- \frac{\kappa_3(t_{n,b})}{6}x^2\phi(x) - \frac{\kappa_4(t_{n,b})}{24}x^3\phi(x) + o(n^{-1}), \tag{3.11}$$

47

and the formal Cornish-Fisher expansions are

$$g(t_{n,b}) = t_{n,b} + (\frac{\kappa_3(t_{n,b})}{6} - \kappa_1(t_{n,b})) + (\frac{\kappa_4(t_{n,b})}{8} - \frac{\kappa_2(t_{n,b}) - 1}{2})t_{n,b}$$

$$- \frac{\kappa_3(t_{n,b})}{6}t_{n,b}^2 - \frac{\kappa_4(t_{n,b})}{24}t_{n,b}^3 + o(n^{-1}), \qquad (3.12)$$

and

$$h(\xi) = \xi + (\kappa_1(t_{n,b}) - \frac{\kappa_3(t_{n,b})}{6}) + (\frac{\kappa_2(t_{n,b}) - 1}{2} - \frac{\kappa_4(t_{n,b})}{8})\xi$$

$$+ \frac{\kappa_3(t_{n,b})}{6}\xi^2 + \frac{\kappa_4(t_{n,b})}{24}\xi^3 + o(n^{-1}). \qquad (3.13)$$

Again we can have the following new uniqueness properties of the Cornish-Fisher expansions for the batch means method. The proofs of the following two propositions are similar to that of Propositions 2.1 and 2.2 and thus omitted.

Proposition 3.5. *Suppose ξ and X_n are random variables, where ξ is $N(0,1)$ and X_n satisfies $\kappa_1(X_n) = l_{n,1} = O(n^{-1/2})$, $\kappa_2(X_n) = 1 + l_{n,2} = 1 + O(n^{-1/2})$, $\kappa_3(X_n) = l_{n,3} = O(n^{-1/2})$, $\kappa_4(X_n) = l_{n,4} = O(n^{-1})$, and $\kappa_i(X_n) = O(n^{-1})$, $i \geq 5$. Then,*

(1) there is a polynomial of degree two, $g_1(X_n) = X_n + a_{n,0} + a_{n,1}X_n + a_{n,2}X_n^2$, such that $a_{n,i} = O(n^{-1/2})$, $i = 0, 1, 2$, and $|\kappa_i(\xi) - \kappa_i(g_1(X_n))| = O(n^{-1})$ for each i, $1 \leq i \leq 3$. Moreover, the coefficients $a_{n,i}$'s are unique up to $O(n^{-1/2})$. Specifically, $a_{n,0} = -l_{n,1} + (1/6)l_{n,3} + o(n^{-1/2})$, $a_{n,1} = -(1/2)l_{n,2} + o(n^{-1/2})$, and $a_{n,2} = -(1/6)l_{n,3} + o(n^{-1/2})$.

(2) If in addition, $\kappa_i(X_n) = o(n^{-1})$, $i \geq 5$, then there is a polynomial of degree three, $g_2(X_n) = X_n + a_{n,0} + a_{n,1}X_n + a_{n,2}X_n^2 + a_{n,3}X_n^3$, such that $a_{n,i} = O(n^{-1/2})$ for each i, $1 \leq i \leq 4$, and $|\kappa_i(\xi) - \kappa_i(g_2(X_n))| = o(n^{-1})$ for each i, $1 \leq i \leq 4$. Moreover, the

48

coefficients $a_{n,i}$'s are unique up to $O(n^{-1})$. Specifically, $a_{n,0} = -l_{n,1} + (1/6)l_{n,3} + o(n^{-1})$, $a_{n,1} = -(1/2)l_{n,2} + (1/8)l_{n,4} + o(n^{-1})$, $a_{n,2} = -(1/6)l_{n,3} + o(n^{-1})$, and $a_{n,3} = -(1/24)l_{n,4} + o(n^{-1})$.

Proposition 3.6. *Suppose ξ and X_n are random variables, where ξ is $N(0,1)$ and X_n satisfies $\kappa_1(X_n) = l_{n,1} = O(n^{-1/2})$, $\kappa_2(X_n) = 1 + l_{n,2} = 1 + O(n^{-1/2})$, $\kappa_3(X_n) = l_{n,3} = O(n^{-1/2})$. Then,*

(1) there is a polynomial of degree two, $h_1(\xi) = \xi + a_{n,0} + a_{n,1}\xi + a_{n,2}\xi^2$, such that $a_{n,i} = O(n^{-1/2})$, $i = 0, 1, 2$, and $|\kappa_i(X_n) - \kappa_i(h_1(\xi))| = O(n^{-1})$ for each i, $1 \leq i \leq 3$. Moreover, the coefficients $a_{n,i}$'s are unique up to $O(n^{-1/2})$. Specifically, $a_{n,0} = l_{n,1} - (1/6)l_{n,3} + o(n^{-1/2})$, $a_{n,1} = (1/2)l_{n,2} + o(n^{-1/2})$, and $a_{n,2} = (1/6)l_{n,3} + o(n^{-1/2})$.

(2) If in addition $\kappa_4(X_n) = l_{n,4} = O(n^{-1})$, then there is a polynomial of degree three, $h_2(\xi) = \xi + a_{n,0} + a_{n,1}\xi + a_{n,2}\xi^2 + a_{n,3}\xi^3$, such taht $a_{n,i} = O(n^{-1/2})$ for each i, $1 \leq i \leq 4$, and $|\kappa_i(X_n) - \kappa_i(h_2(\xi))| = o(n^{-1})$ for each i, $1 \leq i \leq 4$. Moreover, the coefficients $a_{n,i}$'s are unique up to $O(n^{-1})$. Specifically, $a_{n,0} = l_{n,1} - (1/6)l_{n,3} + o(n^{-1})$, $a_{n,1} = (1/2)l_{n,2} - (1/8)l_{n,4} + o(n^{-1})$, $a_{n,2} = (1/6)l_{n,3} + o(n^{-1})$, and $a_{n,3} = (1/24)l_{n,4} + o(n^{-1})$.

3.4. Confidence Intervals Generated from Johnson-Glynn Pivots

In this section, we will generate new confidence interval methods from the Johnson-Glynn pivots and then show that the increases of the lengths of the new confidence intervals generated by the first and second pivots are asymptotically negligible.

First, we discuss how to use the formal Cornish-Fisher expansions (3.12) and (3.13) to produce confidence intervals. In our discussion, we will assume that both $b \sim O(m^{1/2})$ and $n \sim O(m^{1/2})$, namely, b and n approach to infinity roughly at the same rate. We remark that our results in previous two sections are stated in orders of n and b and allow n and b to vary individually. The current choice for $b \sim O(m^{1/2})$ and $n \sim O(m^{1/2})$ here illustrates how confidence intervals can be generated. Other choices of n and b can be be treated in a similar fashion.

Notice that, from Proposition 3.4, the cumulants of $t_{n,b}$ satisfy the assumption of Proposition 3.5. Assume that g_1 and g_2 are the polynomials of degrees two and three, respectively, in Proposition 3.5. Define $T_{n,b} \equiv g_1(t_{n,b})$ and $T_{n,b}^* \equiv g_2(t_{n,b})$. Again, from Propositions 3.4 and 3.5, one can see that $\kappa_1(t_{n,b}) = O(m^{-1/2})$, $\kappa_2(t_{n,b}) = 1 + O(m^{-1/2})$, $\kappa_3(t_{n,b}) = O(m^{-1/2})$, $\kappa_4(t_{n,b}) = O(m^{-1/2})$; $\kappa_1(T_{n,b}) = O(m^{-1})$, $\kappa_2(T_{n,b}) = 1 + O(m^{-1})$, $\kappa_3(T_{n,b}) = O(m^{-1})$, $\kappa_4(T_{n,b}) = O(m^{-1/2})$; $\kappa_1(T_{n,b}^*) = O(m^{-1})$, $\kappa_2(T_{n,b}^*) = 1 + O(m^{-1})$, $\kappa_3(T_{n,b}^*) = O(m^{-1})$, and $\kappa_4(T_{n,b}^*) = O(m^{-1})$. Thus in terms of the distance of the differences between the first four cumulants of two random variables, both $T_{n,b}$ and $T_{n,b}^*$ converge to the standard normal random variable faster than $t_{n,b}$.

In general, we need to estimate the cumulants of $t_{n,b}$. For convenience, we denote, for $i = 1, 2, 3, 4$,

$$\widehat{\kappa_i} \equiv \widehat{\kappa_i(t_{n,b})}, \tag{3.14}$$

as an estimate for the ith cumulant of $t_{n,b}$. Following the terminology in Glynn [12], we define

$$t_{n,b} = (\overline{X} - EX)/(V_{n,b}/n)^{1/2} \tag{3.15}$$

50

as the zeroth order pivot (the usual approach);

$$\widehat{T_{n,b}} = g_1\widehat{(t_{n,b})} \equiv t_{n,b} + (\widehat{\kappa}_3/6 - \widehat{\kappa}_1) - [(\widehat{\kappa}_2 - 1)/2]t_{n,b} - (\widehat{\kappa}_3/6)t_{n,b}^2 \qquad (3.16)$$

as the first order pivot; and

$$\widehat{T_{n,b}^*} = g_2\widehat{(t_{n,b})}$$

$$\equiv t_{n,b} + (\widehat{\kappa}_3/6 - \widehat{\kappa}_1) + [\widehat{\kappa}_4/8 - (\widehat{\kappa}_2 - 1)/2]t_{n,b}$$

$$- (\widehat{\kappa}_3/6)t_{n,b}^2 - (\widehat{\kappa}_4/24)t_{n,b}^3 \qquad (3.17)$$

as the second order pivot. Notice that $t_{n,b}$ is the pivot associated with the traditional batch means method; $\widehat{T_{n,b}}$ is an estimate of $T_{n,b}$, which is the unique polynomial of degree two of $t_{n,b}$ in the sense of Proposition 3.5; and, similarly, $\widehat{T_{n,b}^*}$ is an estimate of $T_{n,b}^*$, which is the corresponding unique polynomial of degree three of t. Thus if the estimators $\widehat{\kappa}_i$'s are reasonably well-behaved, one can expect that both $\widehat{T_{n,b}}$ and $\widehat{T_{n,b}^*}$ converge to the standard normal random variable faster than $t_{n,b}$.

To construct confidence intervals it is computationally more convenient to use the inverted Cornish-Fisher expansions (Hall [12]). For the δ-quantile point, z_δ, of the standard normal distribution function, the $100(1 - 2\delta)\%$ confidence intervals for the three pivots for the batch means method are

$$[\overline{X}_m - h'(z_\delta)(V_{n,b}/n)^{1/2}, \overline{X}_m - h'(-z_\delta)(V_{n,b}/n)^{1/2}], \qquad (3.18)$$

where $h'(z) = z$ for $t_{n,b}$; $h'(z) = z + (\widehat{\kappa}_1 - \widehat{\kappa}_3/6) + [(\widehat{\kappa}_2 - 1)/2]z + (\widehat{\kappa}_3/6)z^2$ for $\widehat{T_{n,b}}$; and $h'(z) = z + (\widehat{\kappa}_1 - \widehat{\kappa}_3/6) + [(\widehat{\kappa}_2 - 1)/2 - (\widehat{\kappa}_4/8)]z + (\widehat{\kappa}_3/6)z^2 + (\widehat{\kappa}_4/24)z^3$ for $\widehat{T_{n,b}^*}$.

It follows from equation (3.18) that the lengths of these $100(1 - 2\delta)\%$ confidence intervals are as follows: $2z_\delta(V_{n,b}/n)^{1/2}$ for the zeroth order pivot $t_{n,b}$; $2z_\delta(1 +$

51

$(\widehat{\kappa_2} - 1)/2)(V_{n,b}/n)^{1/2}$ for the first order pivot $\widehat{T_{n,b}}$; and $[(1+(\widehat{\kappa_2} - 1)/2-(\widehat{\kappa_4}/8))2z_\delta + (\widehat{\kappa_4}/12)z_\delta](V_{n,b}/n)^{1/2}$ for the second order pivot $\widehat{T_{n,b}^*}$. We have shown in Proposition 3.4 that $\kappa_2 - 1 = O(n^{-1}) + O(b^{-1})$ and $\kappa_4 = O(n^{-1})$. Thus if both estimators $\widehat{\kappa_2}$ and $\widehat{\kappa_4}$ are reasonably well-behaved, then one can expect that the increase of the length of the confidence intervals generated by the Johnson-Glynn pivots are asymptotically negligible.

3.5. Computational Efficiency

In this section we will compare the amount of computation required for the traditional batch means method, first and second order Johnson-Glynn pivots for batch means method, and the regenerative method of simulation.

A direct comparison of the first two methods with the regenerative method is difficult. For our purpose, we assume that there are m samples, which are divided into n batches of b samples each, and there are n_0 complete regenerative cycles within these m samples. Note that, by strong law of large number, $n_0 \approx m/E\tau$, where $E\tau$ is the expected length of the regenerative cycles.

We have the following observations.

(1) Generating m samples: for many simulation studies of real world processes, the amount of computation required of generating m samples is of the form $m \cdot c_0 + O(1)$, where c_0 is a constant which depends on the real world system been simulated, implementation of the simulation program, and the computer system used.

(2) Computing batch means: the amount of computation needed is $m \cdot c_1 + O(1)$.

where c_1 does not dependent on m.

(3) Computing sample mean: there are two cases, first, if the sample mean is computed directly from the samples then the amount of computation needed is $m \cdot c_2 + O(1)$; on the other hand, if the sample mean is computed after the batch means are generated then the amount of computation needed is $n \cdot c_2 + O(1)$, where c_2 does not depend on either m or n.

(4) Computing sample variance: the amount of computation needed is $m \cdot c_3 + O(1)$, where c_3, similar to c_0, does not depend on m.

(5) Computing κ_2, κ_3, κ_4 for Johnson-Glynn pivots: the amount of computation needed is $n \cdot c_4 + O(1)$, $n \cdot c_5 + O(1)$, and $n \cdot c_6 + O(1)$, respectively, where c_4, c_5, and c_6 are independent of n but are dependent on the system been simulated, implementation of the simulation program, and the computer system used.

(6) Computing Y_i's and τ_i's for regenerative simulation: the amount of computation needed is $m \cdot c_7 + O(1)$, and $m \cdot c_8 + O(1)$, respectively, where c_7 and c_8 are constants.

(7) Computing $\overline{X}_m$ and $\overline{\tau}$ for regenerative simulation: the amount of computation needed is $n_0 \cdot c_9 + O(1)$, and $n_0 \cdot c_{10} + O(1)$, respectively, where c_9 and c_{10} are constants.

(8) Computing V_{n_0} for regenerative simulation: the amount of computation needed is $n_0 \cdot c_{11} + O(1)$, where c_{11} does not depend on n_0.

(9) It is easy to show that constants $c_1 = c_2 = c_7 = c_8 = c_9 = c_{10}$. Assume that the operations of ith power of a constant will take the same amount of computational effort for $i = 2, 3$, and 4, then we have $c_3 = c_4 = c_5 = c_6$.

We have the following results.

Proposition 3.7. *Under the conditions and constants specified in Observations (1) to (9) above, the amount of computation required for confidence interval methods of simulation output analysis is as follows.*

(1) Traditional i.i.d. method: $m(c_0 + c_1 + c_2) + O(1)$.

(2) Traditional batch means method: $m(c_0 + c_1) + n(c_1 + c_3) + O(1)$.

(3) First order Johnson-Glynn pivot for the batch means method: $m(c_0 + c_1) + n(c_1 + 2c_3) + O(1)$.

(4) Second order Johnson-Glynn pivot for the batch means method: $m(c_0 + c_1) + n(c_1 + 3c_3) + O(1)$.

(5) regenerative method: $m(c_0 + 2c_1) + n_0(2c_1 + c_{11}) + O(1)$.

Proof: (1) Traditional i.i.d. method needs to generate samples, and compute sample mean and sample variance.

(2) Traditional batch means method needs to generate samples, and compute batch means, sample mean, and κ_2.

(3) First order Johnson-Glynn pivot for the batch means method needs to generate samples, and compute batch means, sample mean, κ_2, and κ_3.

(4) Second order Johnson-Glynn pivot for the batch means method needs to generate samples, and compute batch means, sample mean, κ_2, κ_3, and κ_4.

(5) Regenerative method needs to generate samples, and compute Y_i's, τ_i's, $\overline{X}_m$, $\overline{\tau}$, and V_{n_0}. □

It can be seen that the amount of computation required for the Johnson-Glynn pivots for the batch means method is more than that of the traditional batch means

method. However, the relative increments are of $O(b^{-1})$, which is asymptotically negligible as $b \to \infty$, for either cases.

Chapter 4

Numerical Results

In this chapter we report the results of some Monte Carlo studies for the coverage statistics of "normal quantile" confidence intervals based on the Johnson-Glynn pivots $t_{n,b}$, $\widehat{T_{n,b}}$, and $\widehat{T^*_{n,b}}$.

4.1. Notation and Precision

Let us define I_1, I_2, I_3 as follows: I_1 represents the fraction of replications for which the exact value lies to the left of the confidence interval; I_2 represents the fraction that lies in the confidence interval; and I_3 represents the fraction that lies to the right of the confidence interval. Thus I_2 is the usual *coverage fraction*, and I_1 and I_3 are the *one-sided coverage probabilities.*

For each of our examples, we make 2500 independent replications and report empirical coverage fractions I_1, I_2, I_3, sample mean of the length of the confidence interval (SM), sample standard deviation of the length of the confidence interval (SSD), and sample coefficient of variation of the length of the confidence interval (SCV), which is the ratio of sample standard deviation over sample mean (SSD/SM).

Notice that the empirical coverage fractions are essentially the sample means of i.i.d. binomial random variables with suitable parameter p. For 2500 replications and a 90% confidence interval, I_1 and I_3 are the sample means of 2500 i.i.d. binomial

random variables with p approximately equal to 0.05 (5%), which has standard deviation about 0.00436 (0.44%). Similarly, for I_2 the p value of the corresponding i.i.d. binomial random variables is approximately 0.90 (90%), thus the standard deviation of the sample mean is 0.006 (0.6%). These are, we feel, acceptable levels of accuracy for such experiment.

4.2. Examples

Example 4.1. Let $Y_1, \ldots, Y_m$ be i.i.d. random variables with distribution function $P(Y_1 > y) = 1 - \rho e^{-(\mu - \lambda)y}$, $y \leq 0$; 0, otherwise. (We remark parenthetically that this distribution function happens to be the stationary distribution of the waiting times in an M/M/1 queue with mean interarrival time $1/\lambda$ and mean service time $1/\mu$.) See Table 1 for the empirical results of the coverage rate for five different methods of output analysis.

Example 4.2. Let $Y_1, \ldots, Y_m$ be i.i.d. random variables with distribution function $P(Y_1 > y) = e^{-(v+1)}$, $y > -1$; 0, otherwise. Thus each Y_i is a centered-exponential random variable with parameter 1. This example was studied in Efron [11] and Glynn [12]. See Table 2 for the empirical results.

Example 4.3. Let $X_1, \ldots, X_m$ be samples from an autoregressive model of order one such that, for each i, $X_{i+1} = 0.5X_i + \epsilon_{i+1}$, where the residuals ϵ_i's are centered-exponential random variables discussed at Example 4.2. This example was studied in Titus [25]. See Tables 3 and 4 for the empirical results.

Example 4.4. Let $\{W_i : i \geq 1\}$ be a sequence of waiting times in an M/M/1 queue with arrival rate $\lambda = 0.5$ and service rate $\mu = 1$. This example was also

57

studied in Glynn [12] and Titus [25]. To see the effect of initial conditions, we choose several different initial random variable W_1. Let W be the random variable distributed as the stationary waiting times. Those cases we consider include: $W_1 = 0$; $W_1 = W$; $W_1 = EW$; and $W_1 = 2EW$. See Tables 5 to 12 for the empirical results.

4.3. Discussion of Numerical Results

Note that in our numerical examples, essentially in every case there is some improvement in actual coverage fraction I_2 from zeroth order pivot (the traditional batch means method) to the first and second order pivots. As a matter of fact, Table 1 shows that for our Example 1, the first order Johnson-Glynn pivot competes favorably with the method of known variance. In addition, the first and second order pivots tend to balance the one-sided coverage probabilities I_1 and I_3, moving them towards their desired values of 0.05 (5%). This confirms the asymmetry corrections induced by the Johnson-Glynn pivotal transformations. On the other hand, we note that a confidence interval with balanced one-sided coverage probabilities does *not* necessarily have the shortest possible length.

One may notice that in our simulation results, the second order Johnson-Glynn pivot only provides a level of coverage fraction comparable with that of the first order pivot. One reason for this is that in our analysis, we use the formal Cornish-Fisher expansion for $t_{n,b}$ in deriving both pivots $\widehat{T_{n,b}}$ and $\widehat{T^*_{n,b}}$. In this way, $\widehat{T^*_{n,b}}$ is a function of $t_{n,b}$ and its sample cumulants, instead of that of the first order pivot $\widehat{T_{n,b}}$. Hence the $\widehat{T^*_{n,b}}$ we derive has a potential of improvement over $t_{n,b}$, instead of that of $\widehat{T_{n,b}}$. There is, however, another approach in deriving the second

order pivot. If we can derive the formal Cornish-Fisher expansion of $\widehat{T_{n,b}}$, then the second order pivot $\widehat{T^*_{n,b}}$ is then a function of $\widehat{T_{n,b}}$ and its sample cumulants. By the same idea the first order pivot $\widehat{T_{n,b}}$ corrects some effects of estimation errors and the skewness effect associated with $t_{n,b}$, $\widehat{T^*_{n,b}}$ may correct those estimation errors and the skewness effect in the estimation of the first order pivot. Although the second approach is preferable from a theoretical point-of-view, the derivation of an additional pivot from $\widehat{T_{n,b}}$ will be much more difficult.

To examine the effects of nonstationarity and the impact of different initial conditions, we choose several initial random variable W_1 in the simulation of Example 4.4. Let W be the random variable distributed as the stationary waiting times. Those cases we consider include: $W_1 = 0$; $W_1 = W$; $W_1 = EW$; and $W_1 = 2EW$. As we can see empirically from Tables 5 to 12, for the waiting times of an M/M/1 queue, both the nonstationarity and the different initial conditions have only a very insignificant effect on the coverage fraction, one-sided coverage probabilities, and the length of the confidence intervals. Further efforts will be needed to verify this theoretically.

We can also observe that the Johnson-Glynn pivots produce longer confidence intervals on average, and these confidence intervals, in general, are more variable than those of the traditional batch means method. However, as shown in Chapter 3, the increase of length in confidence intervals is asymptotically negligible as n and b increase. Moreover, due to the fact that many confidence intervals produced from the traditional batch means method do have an undercoverage problem, this increase of length seems to be a necessity rather than a liability. The more variable confidence intervals produced by the Johnson-Glynn pivots are also a natural

property associated with these correction methods. Since the correction terms are stochastic, in general more constants need to be estimated and then additional variance is introduced.

Table 1. Coverage Fraction for Example 4.1.

Sample Size (m)	normal	t	first	second	known variance
5	0.6604	0.7620	0.7740	0.7904	0.9312
10	0.7588	0.7912	0.8512	0.8472	0.9180
15	0.8092	0.8260	0.8928	0.8892	0.9152
20	0.8324	0.8464	0.8992	0.8956	0.9080
25	0.8412	0.8532	0.9012	0.8976	0.8988

- 2500 independent replications
- normal: batch means method with normal quantiles
- t: batch means method with t quantiles
- first: first order Johnson-Glynn pivot
- second: second order Johnson-Glynn pivot
- known variance: batch means method with known variance
- $b = 1$, since independently identical observations
- Coverage fraction for the first order pivot competes favorably with that of the batch means method with known variance

Table 2. Coverage Fraction for Centered-Exponential Random Variables.

Sample Size (m)	Pivot	Coverage Fraction			Length of C.I.		
		I_1	I_2	I_3	SM	SSD	SCV
5	normal	0.0408	0.7140	0.2452	1.12	0.65	0.58
	t	0.0188	0.7864	0.1948	1.45	0.84	0.58
	first	0.0152	0.8220	0.1628	1.61	1.01	0.63
	second	0.0160	0.8188	0.1652	1.58	0.98	0.62
10	normal	0.0308	0.7904	0.1788	0.91	0.39	0.42
	t	0.0188	0.8252	0.1560	1.01	0.43	0.42
	first	0.0216	0.8660	0.1124	1.17	0.58	0.49
	second	0.0216	0.8652	0.1132	1.15	0.56	0.48
15	normal	0.0276	0.8276	0.1448	0.78	0.27	0.34
	t	0.0176	0.8488	0.1336	0.84	0.29	0.34
	first	0.0264	0.8884	0.0852	0.96	0.39	0.41
	second	0.0264	0.8880	0.0856	0.94	0.38	0.40
20	normal	0.0240	0.8548	0.1212	0.69	0.21	0.30
	t	0.0192	0.8668	0.1140	0.73	0.22	0.30
	first	0.0244	0.8992	0.0764	0.82	0.29	0.36
	second	0.0264	0.8952	0.0784	0.81	0.28	0.35
25	normal	0.0228	0.8632	0.1144	0.63	0.17	0.27
	t	0.0184	0.8744	0.1072	0.65	0.18	0.27
	first	0.0272	0.8984	0.0744	0.72	0.24	0.32
	second	0.0284	0.8968	0.0748	0.72	0.23	0.32

- 2500 independent replications
- SM: sample mean
- SSD: sample standard deviation
- SCV: sample coefficient of variation
- normal: batch means method with normal quantiles
- t: batch means method with t quantiles
- first: first order Johnson-Glynn pivot
- second: second order Johnson-Glynn pivot
- Both first and second order pivots improve coverage fraction I_2
- One-sided coverage probabilities I_1 and I_3 of the first and second order pivots are more balanced

Table 3. Coverage Fraction for An AR(1) Process.

Batch Size (b)	Pivot	Coverage Fraction			Length of C.I.		
		I_1	I_2	I_3	SM	SSD	SCV
2	normal	0.0788	0.7900	0.1312	0.36	0.055	0.16
	t	0.0768	0.7936	0.1296	0.36	0.056	0.16
	first	0.0932	0.8000	0.1068	0.37	0.063	0.17
	second	0.0936	0.7992	0.1072	0.37	0.062	0.17
4	normal	0.0568	0.8364	0.1068	0.34	0.066	0.19
	t	0.0496	0.8524	0 0980	0.35	0.068	0.19
	first	0.0608	0.8592	0.0800	0.37	0.084	0.23
	second	0.0620	0.8580	0.0800	0.37	0.082	0.22
6	normal	0.0548	0.8460	0.0992	0.33	0.070	0.22
	t	0.0464	0.8664	0.0872	0.34	0.074	0.22
	first	0.0488	0.8764	0.0748	0.36	0.092	0.25
	second	0.0520	0.8728	0.0752	0.36	0.089	0.25
8	normal	0.0464	0.8624	0.1188	0.32	0.077	0.24
	t	0.0384	0.8788	0.0828	0.34	0.082	0.24
	first	0.0408	0.8912	0.0736	0.36	0.10	0.28
	second	0.0424	0.8892	0.0748	0.36	0.099	0.28
10	normal	0.0540	0.8492	0.0968	0.30	0.078	0.26
	t	0.0380	0.8832	0.0788	0.33	0.085	0.26
	first	0.0380	0.8968	0.0652	0.36	0.11	0.30
	second	0.0388	0.8948	0.0664	0.35	0.10	0.29
15	normal	0.0568	0.8312	0.1132	0.29	0.086	0.30
	t	0.0356	0.8812	0.0832	0.33	0.099	0.30
	first	0.0340	0.8944	0.0676	0.36	0.12	0.32
	second	0.0344	0.8920	0.0680	0.35	0.11	0.32

- $X_{i+1} = 0.5\ X_i + \epsilon_{i+1}$
- ϵ_i: centered-exp(1)
- $X_1 = 0$
- sample size: $m = 120$
- 2500 independent replications
- normal: batch means method with normal quantiles
- t: batch means method with t quantiles
- first: first order Johnson-Glynn pivot
- second: second order Johnson-Glynn pivot
- Both first and second order pivots improve coverage fraction I_2
- One-sided coverage probabilities I_1 and I_3 of the first and second order pivots are more balanced

63

Table 4. Coverage Fraction for An AR(1) Process.

Batch Size (b)	Pivot	Coverage Fraction			Length of C.I.		
		I_1	I_2	I_3	SM	SSD	SCV
2	normal	0.0860	0.7884	0.1256	0.26	0.029	0.11
	t	0.0856	0.7904	0.1240	0.26	0.029	0.11
	first	0.1000	0.7924	0.1076	0.26	0.032	0.12
	second	0.1004	0.7920	0.1076	0.26	0.031	0.12
4	normal	0.0568	0.8496	0.0936	0.25	0.034	0.14
	t	0.0536	0.8548	0.0916	0.25	0.034	0.14
	first	0.0644	0.8580	0.0776	0.26	0.039	0.15
	second	0.0648	0.8572	0.0780	0.26	0.038	0.15
6	normal	0.0488	0.8628	0.0884	0.24	0.037	0.16
	t	0.0444	0.8696	0.0860	0.25	0.038	0.16
	first	0.0548	0.8688	0.0764	0.25	0.045	0.18
	second	0.0556	0.8676	0.0768	0.25	0.044	0.17
8	normal	0.0488	0.8664	0.0848	0.23	0.040	0.17
	t	0.0444	0.8768	0.0788	0.24	0.042	0.17
	first	0.0476	0.8832	0.0692	0.25	0.050	0.20
	second	0.0492	0.8816	0.0692	0.25	0.049	0.20
10	normal	0.0468	0.8668	0.0864	0.23	0.041	0.18
	t	0.0388	0.8828	0.0784	0.24	0.043	0.18
	first	0.0440	0.8904	0.0656	0.25	0.051	0.21
	second	0.0452	0.8892	0.0656	0.25	0.050	0.20
15	normal	0.0536	0.8652	0.0812	0.22	0.047	0.22
	t	0.0412	0.8856	0.0732	0.23	0.050	0.22
	first	0.0452	0.8884	0.0664	0.24	0.060	0.24
	second	0.0456	0.8876	0.0668	0.24	0.058	0.24

- $X_{i+1} = 0.5\, X_i + \epsilon_{i+1}$
- ϵ_i: centered-exp(1)
- $X_1 = 0$
- sample size: $m = 240$
- 2500 independent replications
- normal: batch means method with normal quantiles
- t: batch means method with t quantiles
- first: first order Johnson-Glynn pivot
- second: second order Johnson-Glynn pivot
- Both first and second order pivots improve coverage fraction I_2
- One-sided coverage probabilities I_1 and I_3 of the first and second order pivots are more balanced

Table 5. Coverage Fraction for M/M/1 Waiting Times.

Batch Size (b)	Pivot	Coverage Fraction			Length of C.I.		
		I_1	I_2	I_3	SM	SSD	SCV
20	normal	0.0456	0.8104	0.1440	0.45	0.15	0.33
	t	0.0412	0.8200	0.1388	0.46	0.15	0.33
	first	0.0608	0.8408	0.0984	0.52	0.21	0.40
	second	0.0644	0.8368	0.0988	0.51	0.20	0.39
25	normal	0.0392	0.8208	0.1400	0.47	0.16	0.35
	t	0.0376	0.8268	0.1356	0.48	0.17	0.35
	first	0.0520	0.8568	0.0912	0.54	0.28	0.43
	second	0.0540	0.8524	0.0934	0.53	0.22	0.42
40	normal	0.0356	0.8272	0.1372	0.48	0.19	0.39
	t	0.0292	0.8400	0.1308	0.50	0.19	0.39
	first	0.0400	0.8668	0.0932	0.57	0.29	0.50
	second	0.0416	0.8648	0.0936	0.56	0.28	0.49
50	normal	0.0372	0.8212	0.1416	0.48	0.19	0.40
	t	0.0292	0.8408	0.1300	0.51	0.20	0.40
	first	0.0420	0.8608	0.0972	0.58	0.29	0.50
	second	0.0440	0.8584	0.0976	0.57	0.28	0.49
100	normal	0.0388	0.8124	0.1488	0.48	0.21	0.44
	t	0.0260	0.8492	0.1248	0.53	0.23	0.44
	first	0.0284	0.8764	0.0952	0.60	0.32	0.53
	second	0.0296	0.8744	0.0960	0.59	0.31	0.52
125	normal	0.0484	0.7980	0.1536	0.47	0.22	0.47
	t	0.0300	0.8432	0.1268	0.54	0.25	0.47
	first	0.0312	0.8672	0.1016	0.61	0.34	0.56
	second	0.0332	0.8636	0.1032	0.60	0.33	0.55

- M/M/1 waiting times
- $\rho = 0.5$
- $W_1 = W$
- $\{W_i : i \geq 1\}$ is stationary
- sample size: $m = 1000$
- 2500 independent replications
- normal: batch means method with normal quantiles
- t: batch means method with t quantiles
- first: first order Johnson-Glynn pivot
- second: second order Johnson-Glynn pivot
- Both first and second order pivots improve coverage fraction I_2
- One-sided coverage probabilities I_1 and I_3 of the first and second order pivots are more balanced

65

Table 6. Coverage Fraction for M/M/1 Waiting Times.

Batch Size (b)	Pivot	Coverage Fraction			Length of C.I.		
		I_1	I_2	I_3	SM	SSD	SCV
20	normal	0.0476	0.8224	0.1300	0.32	0.082	0.25
	t	0.0464	0.8256	0.1280	0.33	0.083	0.25
	first	0.0656	0.8372	0.0972	0.36	0.10	0.29
	second	0.0680	0.8348	0.0972	0.35	0.10	0.29
25	normal	0.0400	0.8336	0.1264	0.34	0.089	0.27
	t	0.0392	0.8348	0.1260	0.34	0.090	0.27
	first	0.0568	0.8512	0.0920	0.37	0.12	0.32
	second	0.0580	0.8500	0.0920	0.37	0.11	0.31
40	normal	0.0336	0.8436	0.1228	0.35	0.10	0.30
	t	0.0308	0.8488	0.1204	0.36	0.11	0.30
	first	0.0492	0.8652	0.0856	0.39	0.15	0.38
	second	0.0504	0.8624	0.0872	0.39	0.14	0.37
50	normal	0.0300	0.8484	0.1216	0.35	0.11	0.31
	t	0.0292	0.8516	0.1192	0.38	0.11	0.31
	first	0.0448	0.8692	0.0860	0.40	0.16	0.39
	second	0.0468	0.8660	0.0872	0.40	0.15	0.38
100	normal	0.0324	0.8480	0.1196	0.36	0.12	0.34
	t	0.0252	0.8620	0.1128	0.38	0.13	0.34
	first	0.0340	0.8788	0.0872	0.42	0.18	0.42
	second	0.0380	0.8748	0.0872	0.41	0.17	0.41
125	normal	0.0376	0.8328	0.1296	0.35	0.13	0.36
	t	0.0288	0.8516	0.1196	0.38	0.14	0.36
	first	0.0356	0.8720	0.0924	0.42	0.19	0.46
	second	0.0360	0.8712	0.0928	0.42	0.19	0.44

- M/M/1 waiting times
- $\rho = 0.5$
- $W_1 = W$
- $\{W_i : i \geq 1\}$ is stationary
- sample size: $m = 2000$
- 2500 independent replications
- normal: batch means method with normal quantiles
- t: batch means method with t quantiles
- Both first and second order pivots improve coverage fraction I_2
- One-sided coverage probabilities I_1 and I_3 of the first and second order pivots are more balanced

Table 7. Coverage Fraction for M/M/1 Waiting Times.

Batch Size (b)	Pivot	Coverage Fraction			Length of C.I.		
		I_1	I_2	I_3	SM	SSD	SCV
20	normal	0.0428	0.8068	0.1504	0.45	0.15	0.33
	t	0.0396	0.8152	0.1452	0.46	0.15	0.33
	first	0.0592	0.8428	0.0980	0.51	0.20	0.40
	second	0.0624	0.8380	0.0996	0.51	0.20	0.39
25	normal	0.0368	0.8192	0.1440	0.46	0.16	0.35
	t	0.0344	0.8268	0.1388	0.47	0.17	0.35
	first	0.0500	0.8532	0.0968	0.53	0.23	0.43
	second	0.0544	0.8480	0.0976	0.53	0.22	0.42
40	normal	0.0336	0.8216	0.1448	0.48	0.19	0.39
	t	0.0284	0.8336	0.1380	0.50	0.19	0.39
	first	0.0392	0.8648	0.0960	0.57	0.29	0.50
	second	0.0412	0.8620	0.0968	0.54	0.27	0.49
50	normal	0.0336	0.8204	0.1460	0.48	0.19	0.40
	t	0.0288	0.8344	0.1368	0.50	0.20	0.40
	first	0.0392	0.8604	0.1004	0.58	0.29	0.50
	second	0.0412	0.8580	0.1008	0.57	0.28	0.49
100	normal	0.0360	0.8104	0.1536	0.47	0.21	0.44
	t	0.0248	0.8464	0.1288	0.53	0.23	0.44
	first	0.0280	0.8752	0.0968	0.60	0.22	0.53
	second	0.0284	0.8732	0.0984	0.59	0.31	0.52
125	normal	0.0440	0.7968	0.1592	0.47	0.22	0.47
	t	0.0272	0.8412	0.1316	0.54	0.25	0.47
	first	0.0296	0.8632	0.1072	0.60	0.34	0.56
	second	0.0316	0.8592	0.1092	0.60	0.33	0.55

- M/M/1 waiting times
- $\rho = 0.5$
- $W_1 = 0$
- $\{W_i : i \geq 1\}$ is only asymptotically stationary
- sample size: $m = 1000$
- 2500 independent replications
- normal: batch means method with normal quantiles
- t: batch means method with t quantiles
- first: first order Johnson-Glynn pivot
- second: second order Johnson-Glynn pivot
- Both first and second order pivots improve coverage fraction I_2
- One-sided coverage probabilities I_1 and I_3 of the first and second order pivots are more balanced

Table 8. Coverage Fraction for M/M/1 Waiting Times.

Batch Size (b)	Pivot	Coverage Fraction			Length of C.I.		
		I_1	I_2	I_3	SM	SSD	SCV
20	normal	0.0456	0.8216	0.1329	0.33	0.082	0.25
	t	0.0452	0.8228	0.1320	0.33	0.083	0.25
	first	0.0640	0.8372	0.0988	0.36	0.10	0.29
	second	0.0660	0.8340	0.1000	0.35	0.10	0.29
25	normal	0.0376	0.8336	0.1288	0.34	0.089	0.27
	t	0.0368	0.8356	0.1276	0.34	0.090	0.27
	first	0.0560	0.8520	0.0920	0.37	0.12	0.32
	second	0.0572	0.8500	0.0928	0.37	0.11	0.31
40	normal	0.0332	0.8416	0.1252	0.35	0.10	0.30
	t	0.0308	0.8472	0.1220	0.36	0.11	0.30
	first	0.0464	0.8672	0.0864	0.39	0.15	0.38
	second	0.0476	0.8652	0.0872	0.39	0.14	0.37
50	normal	0.0304	0.8456	0.1240	0.35	0.11	0.31
	t	0.0272	0.8516	0.1212	0.36	0.11	0.31
	first	0.0420	0.8680	0.0868	0.40	0.16	0.39
	second	0.0452	0.8840	0.0884	0.40	0.15	0.38
100	normal	0.0348	0.8472	0.1236	0.36	0.12	0.34
	t	0.0228	0.8628	0.1144	0.37	0.13	0.34
	first	0.0316	0.8760	0.0924	0.42	0.18	0.42
	second	0.0352	0.8724	0.0924	0.41	0.17	0.41
125	normal	0.0348	0.8324	0.1328	0.35	0.13	0.36
	t	0.0268	0.8520	0.1212	0.38	0.13	0.36
	first	0.0336	0.8704	0.0960	0.42	0.19	0.46
	second	0.0336	0.8700	0.0964	0.42	0.18	0.44

- M/M/1 waiting times
- $\rho = 0.5$
- $W_1 = 0$
- $\{W_i : i \geq 1\}$ is only asymptotically stationary
- sample size: $m = 2000$
- 2500 independent replications
- normal: batch means method with normal quantiles
- t: batch means method with t quantiles
- first: first order Johnson-Glynn pivot
- second: second order Johnson-Glynn pivot
- Both first and second order pivots improve coverage fraction I_2
- One-sided coverage probabilities I_1 and I_3 of the first and second order pivots are more balanced

Table 9. Coverage Fraction for M/M/1 Waiting Times.

Batch Size (b)	Pivot	Coverage Fraction			Length of C.I.		
		I_1	I_2	I_3	SM	SSD	SCV
20	normal	0.0440	0.8092	0.1468	0.45	0.15	0.33
	t	0.0400	0.8212	0.1388	0.46	0.15	0.33
	first	0.0600	0.8440	0.0960	0.51	0.20	0.40
	second	0.0636	0.8392	0.0972	0.51	0.20	0.39
25	normal	0.0380	0.8220	0.1400	0.46	0.16	0.35
	t	0.0348	0.8288	0.1364	0.47	0.17	0.35
	first	0.0524	0.8516	0.0960	0.53	0.23	0.42
	second	0.0560	0.8472	0.0968	0.53	0.22	0.41
40	normal	0.0344	0.8256	0.1400	0.48	0.19	0.39
	t	0.0292	0.8348	0.1360	0.50	0.19	0.39
	first	0.0396	0.8672	0.0932	0.57	0.28	0.50
	second	0.0420	0.8640	0.0940	0.56	0.27	0.49
50	normal	0.0344	0.8232	0.1424	0.48	0.19	0.40
	t	0.0288	0.8388	0.1324	0.50	0.20	0.40
	first	0.0404	0.8608	0.0988	0.58	0.29	0.50
	second	0.0428	0.8576	0.0996	0.57	0.28	0.49
100	normal	0.0388	0.8108	0.1504	0.47	0.21	0.44
	t	0.0252	0.8492	0.1252	0.53	0.23	0.44
	first	0.0284	0.8780	0.0936	0.60	0.32	0.53
	second	0.0296	0.8748	0.0956	0.59	0.31	0.52
125	normal	0.0456	0.7972	0.1572	0.47	0.22	0.47
	t	0.0280	0.8436	0.1284	0.54	0.25	0.47
	first	0.0312	0.8652	0.1036	0.60	0.34	0.56
	second	0.0324	0.8628	0.1048	0.60	0.33	0.55

- M/M/1 waiting times
- $\rho = 0.5$
- $W_1 = EW$
- $\{W_i : i \geq 1\}$ is only asymptotically stationary
- sample size: $m = 1000$
- 2500 independent replications
- normal: batch means method with normal quantiles
- t: batch means method with t quantiles
- first: first order Johnson-Glynn pivot
- second: second order Johnson-Glynn pivot
- Both first and second order pivots improve coverage fraction I_2
- One-sided coverage probabilities I_1 and I_3 of the first and second order pivots are more balanced

Table 10. Coverage Fraction for M/M/1 Waiting Times.

Batch Size (b)	Pivot	Coverage Fraction			Length of C.I.		
		I_1	I_2	I_3	SM	SSD	SCV
20	normal	0.0468	0.8212	0.1320	0.33	0.082	0.25
	t	0.0452	0.8240	0.1308	0.33	0.083	0.25
	first	0.0648	0.8388	0.0964	0.36	0.10	0.29
	second	0.0664	0.8364	0.0972	0.35	0.10	0.29
25	normal	0.0388	0.8332	0.1280	0.34	0.089	0.26
	t	0.0376	0.8352	0.1272	0.34	0.090	0.26
	first	0.0572	0.8516	0.0912	0.37	0.12	0.32
	second	0.0580	0.8500	0.0920	0.37	0.11	0.31
40	normal	0.0332	0.8436	0.1232	0.35	0.10	0.30
	t	0.0316	0.8488	0.1196	0.36	0.11	0.30
	first	0.0476	0.8668	0.0856	0.39	0.15	0.38
	second	0.0484	0.8652	0.0864	0.39	0.14	0.37
50	normal	0.0308	0.8456	0.1236	0.35	0.11	0.31
	t	0.0276	0.8524	0.1200	0.36	0.11	0.31
	first	0.0428	0.8712	0.0860	0.40	0.16	0.39
	second	0.0464	0.8656	0.0880	0.40	0.15	0.38
100	normal	0.0304	0.8476	0.1220	0.36	0.12	0.34
	t	0.0228	0.8644	0.1128	0.37	0.13	0.34
	first	0.0336	0.8760	0.0904	0.42	0.18	0.42
	second	0.0360	0.8728	0.0912	0.41	0.17	0.41
125	normal	0.0352	0.8332	0.1316	0.35	0.13	0.36
	t	0.0272	0.8524	0.1204	0.38	0.13	0.36
	first	0.0336	0.8720	0.0944	0.42	0.19	0.46
	second	0.0340	0.8708	0.0952	0.42	0.18	0.44

- M/M/1 waiting times
- $\rho = 0.5$
- $W_1 = EW$
- $\{W_i : i \geq 1\}$ is only asymptotically stationary
- sample size: $m = 2000$
- 2500 independent replications
- normal: batch means method with normal quantiles
- t: batch means method with t quantiles
- Both first and second order pivots improve coverage fraction I_2
- One-sided coverage probabilities I_1 and I_3 of the first and second order pivots are more balanced

70

Table 11. Coverage Fraction for M/M/1 Waiting Times.

Batch Size (b)	Pivot	Coverage Fraction			Length of C.I.		
		I_1	I_2	I_3	SM	SSD	SCV
20	normal	0.0452	0.8164	0.1384	0.45	0.15	0.33
	t	0.0412	0.8248	0.1340	0.46	0.15	0.33
	first	0.0640	0.8444	0.0920	0.52	0.20	0.40
	second	0.0672	0.8400	0.0916	0.51	0.20	0.39
25	normal	0.0392	0.8244	0.1364	0.46	0.16	0.35
	t	0.0356	0.8316	0.1328	0.48	0.17	0.35
	first	0.0540	0.8544	0.0916	0.54	0.23	0.42
	second	0.0572	0.8492	0.0936	0.53	0.22	0.41
40	normal	0.0356	0.8276	0.1368	0.48	0.19	0.39
	t	0.0308	0.8396	0.1296	0.50	0.19	0.39
	first	0.0420	0.8672	0.0908	0.57	0.29	0.50
	second	0.0440	0.8644	0.0916	0.56	0.27	0.49
50	normal	0.0384	0.8240	0.1376	0.48	0.19	0.40
	t	0.0300	0.8424	0.1276	0.50	0.20	0.40
	first	0.0428	0.8640	0.0932	0.58	0.29	0.50
	second	0.0440	0.8612	0.0948	0.57	0.28	0.49
100	normal	0.0400	0.8152	0.1448	0.47	0.21	0.44
	t	0.0260	0.8536	0.1204	0.53	0.23	0.44
	first	0.0292	0.8824	0.0884	0.60	0.32	0.53
	second	0.0304	0.8800	0.0896	0.59	0.31	0.52
125	normal	0.0476	0.7996	0.1528	0.47	0.22	0.47
	t	0.0308	0.8468	0.1224	0.54	0.25	0.47
	first	0.0324	0.8688	0.0988	0.61	0.34	0.56
	second	0.0324	0.8672	0.1004	0.60	0.33	0.55

- M/M/1 waiting times
- $\rho = 0.5$
- $W_1 = 2\,EW$
- $\{W_i : i \geq 1\}$ is only asymptotically stationary
- sample size: $m = 1000$
- 2500 independent replications
- normal: batch means method with normal quantiles
- t: batch means method with t quantiles
- first: first order Johnson-Glynn pivot
- second: second order Johnson-Glynn pivot
- Both first and second order pivots improve coverage fraction I_2
- One-sided coverage probabilities I_1 and I_3 of the first and second order pivots are more balanced

Table 12. Coverage Fraction for M/M/1 Waiting Times.

Batch Size (b)	Pivot	Coverage Fraction			Length of C.I.		
		I_1	I_2	I_3	SM	SSD	SCV
20	normal	0.0488	0.8204	0.1308	0.33	0.082	0.25
	t	0.0472	0.8240	0.1288	0.33	0.083	0.25
	first	0.0664	0.8396	0.0940	0.36	0.10	0.29
	second	0.0704	0.8348	0.0948	0.35	0.10	0.29
25	normal	0.0408	0.8344	0.1248	0.34	0.089	0.27
	t	0.0380	0.8388	0.1232	0.34	0.090	0.27
	first	0.0580	0.8540	0.0880	0.37	0.12	0.32
	second	0.0592	0.8512	0.0896	0.37	0.11	0.31
40	normal	0.0352	0.8436	0.1212	0.35	0.10	0.30
	t	0.0328	0.8496	0.1176	0.36	0.11	0.30
	first	0.0484	0.8684	0.0832	0.39	0.15	0.38
	second	0.0496	0.8668	0.0836	0.39	0.14	0.37
50	normal	0.0320	0.8476	0.1204	0.35	0.11	0.31
	t	0.0276	0.8568	0.1156	0.36	0.11	0.31
	first	0.0440	0.8712	0.0844	0.40	0.16	0.39
	second	0.0472	0.8672	0.0856	0.40	0.15	0.38
100	normal	0.0316	0.8500	0.1184	0.36	0.12	0.34
	t	0.0236	0.8664	0.1100	0.37	0.13	0.34
	first	0.0352	0.8784	0.0864	0.42	0.18	0.42
	second	0.0368	0.8752	0.0880	0.41	0.17	0.41
125	normal	0.0364	0.8348	0.1288	0.35	0.13	0.36
	t	0.0276	0.8564	0.1160	0.38	0.13	0.36
	first	0.0348	0.8732	0.0920	0.42	0.19	0.46
	second	0.0356	0.8724	0.0920	0.42	0.18	0.44

- M/M/1 waiting times
- $\rho = 0.5$
- $W_1 = 2\,EW$
- $\{W_i : i \geq 1\}$ is only asymptotically stationary
- sample size: $m = 2000$
- 2500 independent replications
- normal: batch means method with normal quantiles
- t: batch means method with t quantiles
- first: first order Johnson-Glynn pivot
- second: second order Johnson-Glynn pivot
- Both first and second order pivots improve coverage fraction I_2
- One-sided coverage probabilities I_1 and I_3 of the first and second order pivots are more balanced

Chapter 5

Conclusion

As the use of simulation becomes more popular, the need for a method of simulation output analysis that can be applied to a large class of stochastic processes becomes more important.

In this dissertation, we have derived a uniqueness property of Cornish-Fisher expansions, the formal Edgeworth expansion for the batch means method, Johnson-Glynn pivots and the associated confidence intervals for the batch means method, and detailed procedures of implementation.

Johnson-Glynn pivotal transformations have provided a new way of generating confidence intervals. In applying this approach to the batch means method, they appear to behave well empirically and seem to be a robust procedure for the examples in Chapter 4. To verify the general applicability of this method, many more examples should be run.

However, there are three possible drawbacks from a practical point-of-view: more computation time needed, longer confidence intervals on average, and more variable intervals.

On the other hand, as shown in Chapter 3, the increase of computing time is relatively small, and the increase of length in confidence interval is asymptotically negligible as n and b increase. Moreover, due to the fact that many confidence intervals do have an undercoverage problem, this increase of length seems to be a necessity rather than a liability.

We would like to point out some potential areas for future research and development. First, although the assumptions we make are reasonable, we may want to relax the assumptions we made on stationarity, either by including the initial condition or considering asymptotically stationary processes. Second, it is possible that the Johnson-Glynn pivots can be applied to overlapping batch means method. Another interesting possibility is to apply the same procedure to the ratio estimator of a weakly regenerative process. Finally, we may want to run many more numerical examples to verify the general applicability of this method.

Appendix

Proof of Proposition 2.1: (1) Suppose $g_1(X_n) = X_n + a_{n,0} + a_{n,1}X_n + a_{n,2}X_n^2$, where $a_i = O(n^{-1/2})$, $i = 0, 1$, and 2. Then $\kappa_1(g_1(X_n)) = EX_n + a_{n,0} + a_{n,1}EX_n + a_{n,2}EX_n^2 = l_{n,1} + a_{n,0} + a_{n,2} + O(n^{-1})$, $\kappa_2(g_1(X_n)) = 1 + 2a_{n,1} + l_{n,2} + O(n^{-1})$, $\kappa_3(g_1(X_n)) = l_{n,3} + 6a_{n,2} + O(n^{-1})$. Notice that $\kappa_1(\xi) = 0$, $\kappa_2(\xi) = 1$, and $\kappa_3(\xi) = 0$. If we need $|\kappa_i(\xi) - \kappa_i(g_1(X_n))| = O(n^{-1})$ for each i, $1 \leq i \leq 3$, the unique solution of a_i's in $g_1(X_n)$, up to order $O(n^{-1/2})$ is such that $a_{n,0} = -l_{n,1} + l_{n,3}/6 + O(n^{-1})$, $a_{n,1} = O(n^{-1})$, and $a_{n,2} = -l_{n,3}/6 + O(n^{-1})$.

(2) Suppose $g_2(X_n) = X_n + a_{n,0} + a_{n,1}X_n + a_{n,2}X_n^2 + a_{n,3}X_n^3$, and $a_i = O(n^{-1/2})$ for $i = 0, 1, 2$, and 3. Then $\kappa_1(g_2(X_n)) = l_{n,1} + a_{n,0} + a_{n,1}l_{n,1} + a_{n,2} + a_{n,3}l_{n,3} + o(n^{-1})$, $\kappa_2(g_2(X_n)) = 1 + 2a_{n,1} + a_{n,1}^2 + l_{n,2} + 2a_{n,2}^2 + 15a_{n,3}^2 + 2a_{n,2}l_{n,3} + 4a_{n,2}l_{n,1} + 6a_{n,3} + 6a_{n,1}a_{n,3} + o(n^{-1})$, $\kappa_3(g_2(X_n)) = l_{n,3} + 3a_{n,1}l_{n,3} + 6a_{n,2} + 12a_{n,1}a_{n,2} + 27a_{n,3}l_{n,3} + 18a_{n,3}l_{n,1} + o(n^{-1})$, $\kappa_4(g_2(X_n)) = l_{n,4} + 24a_{n,2}l_{n,3} + 24a_{n,3} + 72a_{n,1}a_{n,3} + 48a_{n,2}^2 + 432a_{n,3}^2 + o(n^{-1})$. If we need $|\kappa_i(\xi) - \kappa_i(g_2(X_n))| = o(n^{-1})$ for each i, $1 \leq i \leq 4$, the unique solution of a_i's in $g_2(X_n)$, up to order $O(n^{-1})$ is such that $a_{n,0} = -l_{n,1} + l_{n,3}/6 + o(n^{-1})$, $a_{n,1} = -l_{n,2}/2 + l_{n,1}l_{n,3}/3 + l_{n,4}/8 - (7/36)l_{n,2}^2 + o(n^{-1})$, $a_{n,2} = -l_{n,3}/6 + o(n^{-1})$, and $a_{n,3} = -l_{n,4}/24 + l_{n,3}^2/9 + o(n^{-1})$. $\square$

Proof of Proposition 2.2: (1) Suppose $h_1(\xi) = \xi + a_{n,0} + a_{n,1}\xi + a_{n,2}\xi^2$, where $a_i = O(n^{-1/2})$, $i = 0, 1$, and 2. Then $\kappa_1(h_1(\xi)) = a_{n,0} + a_{n,2}$, $\kappa_2(h_1(\xi)) = 1 + 2a_{n,1} + O(n^{-1})$. $\kappa_3(h_1(\xi)) = 6a_{n,2} + O(n^{-1})$. Notice that $\kappa_1(X_n) = 0$, $\kappa_2(X_n) = 1$.

and $\kappa_3(X_n) = l_{n,3} = O(n^{-1/2})$. If we need $|\kappa_i(X_n) - \kappa_i(h_1(\xi))| = O(n^{-1})$ for each i, $1 \le i \le 3$, the unique solution of a_i's in $h_1(\xi)$, up to order $O(n^{-1/2})$ is such that $a_{n,0} = l_{n,1} - l_{n,3}/6 + O(n^{-1})$, $a_{n,1} = O(n^{-1})$, and $a_{n,2} = l_{n,3}/6 + O(n^{-1})$.

(2) Suppose $h_2(\xi) = \xi + a_{n,0} + a_{n,1}\xi + a_{n,2}\xi^2 + a_{n,3}\xi^3$, and $a_i = O(n^{-1/2})$ for $i = 0, 1$, 2, and 3. Then $\kappa_1(h_2(\xi)) = a_{n,0} + a_{n,2}$, $\kappa_2(h_2(\xi)) = 1 + 2a_{n,1} + a_{n,1}^2 + 2a_{n,2}^2 + 15a_{n,3}^2 + 6a_{n,3} + 6a_{n,1}a_{n,3}$, $\kappa_3(h_2(\xi)) = 6a_{n,2} + 12a_{n,1}a_{n,2} + 72a_{n,2}a_{n,3} + o(n^{-1})$, $\kappa_4(h_2(\xi)) = 24a_{n,3} + 48a_{n,1}a_{n,3} + 48a_{n,2}^2 + 432a_{n,3}^2 + o(n^{-1})$. If we need $|\kappa_i(X_n) - \kappa_i(h_2(\xi))| = o(n^{-1})$ for each i, $1 \le i \le 4$, the unique solution of a_i's in $h_2(\xi)$, up to order $O(n^{-1})$ is such that $a_{n,0} = l_{n,1} - l_{n,3}/6 + o(n^{-1})$, $a_{n,1} = l_{n,2}/2 - l_{n,4}/8 + (5/36)l_{n,3}^2 + o(n^{-1})$. $a_{n,2} = l_{n,3}/6 + o(n^{-1})$, and $a_{n,3} = l_{n,4}/24 - l_{n,3}^2/18 + o(n^{-1})$. $\qquad\square$

Proof of Proposition 3.1: Results (1), (2), and (4) follow from the definition of $\{Y_i : i \ge 1\}$. For (3), we have

$$h(\beta_1) + bh(\beta_2) + bh(\beta_3) + \cdots$$

$$\le h(\alpha_1) + [h(\alpha_2) + \cdots + h(\alpha_{b+1})] + [h(\alpha_{b+2}) + \cdots + h(\alpha_{2b+1})] + \cdots$$

$$= \sum_{i=1}^{\infty} h(\alpha_i)$$

$$= [h(\alpha_1) + \cdots + h(\alpha_b)] + [h(\alpha_{b+1}) + \cdots + h(\alpha_{2b})] + \cdots$$

$$\le bh(\beta_1) + bh(\beta_2) + bh(\beta_3) + \cdots,$$

where the first inequality comes from $\beta_i = \alpha_{(i-1)b+1} \le \alpha_{(i-1)b+1-j}$, for all $j \ge 0$ and the monotonicity of h. Similarly, the last inequality is a consequence of $\beta_i \ge \alpha_{(i-1)b+j}$, for all $j \ge 1$ and the monotonicity of h. The result then follows. $\qquad\square$

Two more prositions remain to be proved. These will require a number of lemmas which are given below.

Lemma A.1. *Suppose* $Y \in \sigma(X_1, \ldots, X_k)$ *and* $Z \in \sigma(X_{k+n}, X_{k+n+1}, \ldots)$.

(1) If $|Y|$ *is bounded by* C *and* $|Z|$ *is bounded by* D, *then*

$$|\mathrm{cum}(Y, Z)| \leq 4CD\alpha_n.$$

(2) If $E[Y^4] \leq C$ *and* $E[Z^4] \leq D$, *then*

$$|\mathrm{cum}(Y, Z)| \leq 4(1 + C + D)\alpha_n^{1/2}.$$

(3) If $E[Y^4] \leq C$ *and* $E[Z^4] \leq D$, *then for positive numbers* M *and* N,

$$|\mathrm{cum}(Y, Z)| \leq 4MN\alpha_n + 4DNM^{-3} + 4CMN^{-3} + C^{1/2}D^{1/2}M^{-1}N^{-1}.$$

Proof: (1) See Lemma 2 of Billingsley [5], p. 317.

(2) See Lemma 3 of Billingsley [5], p. 317.

(3) Let A be a set and define the *indicator function* I_A take value 1 on A and 0 elsewhere. Let $Y_0 = YI_{|Y| \leq N}$, $Y_1 = YI_{|Y| > N}$, $Z_0 = YI_{|Z| \leq M}$, and $Z_1 = ZI_{|Z| > M}$. Then $|E[Y_0 Z_0] - E[Y_0]E[Z_0]| \leq 4MN\alpha_n$, $|E[Y_0 Z_1] - E[Y_0]E[Z_1]| \leq 4DM^{-3}N$, $|E[Y_1 Z_0] - E[Y_1]E[Z_0]| \leq 4CMN^{-3}$, and $|E[Y_1 Z_1] - E[Y_1]E[Z_1]| \leq C^{1/2}D^{1/2}M^{-1}N^{-1}$. The result follows from $|\mathrm{cum}(Y, Z)| = |E[YZ] - E[Y]E[Z]| \leq |E[Y_0 Z_0] - E[Y_0]E[Z_0]| + |E[Y_0 Z_1] - E[Y_0]E[Z_1]| + |E[Y_1 Z_0] - E[Y_1]E[Z_0]| + |E[Y_1 Z_1] - E[Y_1]E[Z_1]|.$ □

Lemma A.2. *Suppose* $W \in \sigma(X_1, \ldots, X_k)$, $Y \in \sigma(X_{k+n}, \ldots, X_l)$, *and* $Z \in \sigma(X_{l+m}, X_{l+m+1}, \ldots)$. *Let* $\alpha_{m,n} \equiv \min\{\alpha_m, \alpha_n, \alpha_{m+n}\}$. *Then*

(1) If $|W|$ *is bounded by* B, $|Y|$ *is bounded by* C, *and* $|Z|$ *is bounded by* D, *then*

$$|\mathrm{cum}(W, Y, Z)| \leq 8BCD\alpha_{m,n}.$$

(2) If $E[|W^4|] \leq B$, $E[|Y^4|] \leq C$, *and* $E[|Z^4|] \leq D$, *then*

$$|\mathrm{cum}(W, Y, Z)| \leq 8LMN\alpha_{m,n}$$

$$+ 8MNBL^{-3} + 8LNCM^{-3} + 8LMDN^{-3}$$

$$+ 2LC^{1/2}D^{1/2}M^{-1}N^{-1} + 2MB^{1/2}D^{1/2}L^{-1}N^{-1}$$

$$+ 2NB^{1/2}C^{1/2}L^{-1}M^{-1} + B^{1/3}C^{1/3}D^{1/3}L^{-1/3}M^{-1/3}N^{-1/3}.$$

(3) If $E[|W^4|] \leq B$, $E[|Y^4|] \leq C$, and $E[|Z^4|] \leq D$, then

$$|\text{cum}(W, Y, Z)| \leq c\alpha_{m,n}^{1/4},$$

where c is a constant.

Proof: (1) Notice that $|\text{cum}(W, Y, Z)| = |E[W - E(W)][Y - E(Y)][Z - E(Z)]| \leq 2B|E[Y - E(Y)][Z - E(Z)]| \leq 8BCD\alpha_n$. Similarly, $|\text{cum}(W, Y, Z)| \leq 8BCD\alpha_m$ and $|\text{cum}(W, Y, Z)| \leq 8BCD\alpha_{m+n}$.

(2) Similar to the proof of Lemma A.1(3).

(3) The desired result can be obtained by using (2) and let $L = M = N \sim \alpha_{m,n}^{-1/4}$.

$\square$

Lemma A.3. Suppose $\sum_{n=-\infty}^{\infty} \alpha_n^{1/2} < \infty$.

(1) If $EY_0^4 \leq Cb^{-2}$ for some constant C, then $\sum_{n=-\infty}^{\infty} |\text{Cov}(Y_0, Y_n)| = O(b^{-1})$.

(2) If $EY_0^8 \leq Db^{-4}$ for some constant D, then $\sum_{n=-\infty}^{\infty} |\text{Cov}(Y_0, Y_n^2)| = O(b^{-2})$.

(3) If $EY_0^8 \leq Db^{-4}$ for some constant D, then $\sum_{n=-\infty}^{\infty} |\text{Cov}(Y_0^2, Y_n^2)| = O(b^{-2})$.

Proof: (1) From Lemma A.1, $|\text{cum}(Y_0, Y_i)| \leq 4N^2\beta_i + 4Cb^{-4}N^{-2} + 4Cb^{-2}N^{-2} + Cb^{-3}N^{-2}$. Let $N = b^{-1/2}\beta_i^{-1/4}$, the above ineaqality can be reduced to $|\text{cum}(Y_0, Y_i)| \leq cb^{-1}\beta_i^{1/2}$ so that $\sum_{i=-\infty}^{\infty} |\text{cum}(Y_0, Y_i)| \leq cb^{-1}\sum_{i=-\infty}^{\infty}\beta_i^{1/2} \leq cb^{-1}\sum_{i=-\infty}^{\infty}\alpha_i^{1/2} = O(b^{-1})$.

(2) From Liapounov inequality, $(E|Y_0|^4)^{1/4} \leq (E|Y_0|^8)^{1/8}$, it follows that $EY_0^4 \leq$

Cb^{-2} where $C = D^{1/2}$. From Lemma A.1, $|\text{cum}(Y_0, Y_i^2)| \leq 4MN\beta_i + 4Db^{-4}M^{-3}N +$

$4Cb^{-2}MN^{-3} + C^{1/2}D^{1/2}b^{-3}M^{-1}N^{-1}$. Let $N = b^{-1/2}\beta_i^{-1/4}$ and $M = b^{-1}\beta_i^{-1/4}$, the

above ineaqality can be reduced to $|\text{cum}(Y_0, Y_i^2)| \leq cb^{-3/2}\beta_i^{1/2}$ so that

$\sum_{i=-\infty}^{\infty} |\text{cum}(Y_0, Y_i^2)| \leq cb^{-3/2} \sum_{i=-\infty}^{\infty} \beta_i^{1/2} \leq cb^{-3/2} \sum_{i=-\infty}^{\infty} \alpha_i^{1/2} = O(b^{-3/2})$. But

there is no $O(b^{-3/2})$ terms on the left hand side, it follows that $\sum_{n=-\infty}^{\infty} |\text{Cov}(Y_0, Y_n^2)|$

is at most $O(b^{-2})$.

(3) From Lemma A.1, $|\text{cum}(Y_0^2, Y_i^2)| \leq 4N^2\beta_i + 9Db^{-4}N^{-2}$. Let $N = b^{-1}\beta_i^{-1/4}$

then $|\text{cum}(Y_0^2, Y_i^2)| \leq cb^{-2} \sum_{i=-\infty}^{\infty} \beta_i^{1/2}$. The result now follows. $\qquad \square$

Lemma A.4. *Assume* $\sum_{m,n=-\infty}^{\infty} \alpha_{m,n}^{1/4} < \infty$.

(1) If $EY_0^4 \leq Cb^{-2}$ *for some constants* C, *then* $\sum_{m,n=-\infty}^{\infty} |\text{cum}(Y_0, Y_m, Y_n)| = O(b^{-2})$.

(2) If $EY_0^8 \leq Db^{-4}$ *for some constant* D, *then* $\sum_{m,n=-\infty}^{\infty} |\text{cum}(Y_0, Y_m, Y_n^2)| = O(b^{-2})$.

(3) If $EY_0^8 \leq Db^{-4}$ *for some constant* D, *then* $\sum_{m,n=-\infty}^{\infty} |\text{cum}(Y_0, Y_m^2, Y_n^2)| = O(b^{-3})$.

(4) If $EY_0^8 \leq Db^{-4}$ *for some constants* D, *then* $\sum_{m,n=-\infty}^{\infty} |\text{cum}(Y_0^2, Y_m^2, Y_n^2)| = O(b^{-3})$.

Proof: (1) Let $L = b^{-1/2}\beta_{m,n}^{-1/4}$, $M = b^{-1/2}\beta_{m,n}^{-1/4}$, and $N = b^{-1/2}\beta_{m,n}^{-1/4}$ and then

from Lemma A.2, $|\text{cum}(Y_0, Y_m, Y_n^2)| \leq cb^{-2}\beta_{m,n}^{1/4}$, where c is a constant, so that the

result follows.

(2) Similar to (1) except choosing $L = b^{-1/2}\beta_{m,n}^{-1/4}$, $M = b^{-1/2}\beta_{m,n}^{-1/4}$, and $N = b^{-1}\beta_{m,n}^{-1/4}$.

(3) Similar to (1) except choosing $L = b^{-1/2}\beta_{m,n}^{-1/4}$, $M = b^{-1}\beta_{m,n}^{-1/4}$, and $N = b^{-1}\beta_{m,n}^{-1/4}$.

(4) Similar to (1) except choosing $L = b^{-1}\beta_{m,n}^{-1/4}$, $M = b^{-1}\beta_{m,n}^{-1/4}$, and $N = b^{-1}\beta_{m,n}^{-1/4}$.

$\square$

Remark A.5. *Basically, in Lemmas A.1 through A.4, we have demonstrated that an infinite sum of cumulants such as*

$$\sum |\mathrm{cum}[Y_0^{i_0}, Y_{n_1}^{i_1}, \ldots, Y_{n_k}^{i_k}]|$$

can be calculated by the following approach. We first show that each summand is bounded by, say, $cb^{-i}h(\beta, n_1, \ldots, n_k)$, where c is a constant, b is the batch size, and $h(\cdot)$ is a real value function. However, by suitable assumptions of bounded moments and asymptotic properties of mixing constants, the infinite sum

$$\sum h(\beta, n_1, \ldots, n_k)$$

is finite so that $\sum |\mathrm{cum}[Y_0^{i_0}, Y_{n_1}^{i_1}, \ldots, Y_{n_k}^{i_k}]|$ is of order $O(b^{-i})$.

Corollary A.6 (Cumulants of $\overline{X}_m^2$). *Assume that $\{X_i : i \geq 1\}$ is a discrete time, strictly stationary stochastic process with zero mean.*

(1) If $E|X_1|^4 < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-4-\epsilon})$ for some $\epsilon > 0$, then $\kappa_1(\overline{X}_m^2) = O(m^{-1})$.

(2) If $E|X_1|^{12} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-8-\epsilon})$ for some $\epsilon > 0$. then $\kappa_2(\overline{X}_m^2) = O(m^{-2})$.

(3) If $E|X_1|^{20} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-12-\epsilon})$ for some $\epsilon > 0$. then $\kappa_3(\overline{X}_m^2) = O(m^{-3})$.

(4) If $E|X_1|^{28} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-16-\epsilon})$ for some $\epsilon > 0$, then $\kappa_4(\overline{X}_m^2) = O(m^{-4})$.

Proof: All four results are special cases of Lemma 2.6. $\qquad\square$

Corollary A.7 (Mixed cumulants of $\overline{X}_m$ and $\overline{X}_m^2$). *Assume that $\{X_i : i \geq 1\}$ is a discrete time, strictly stationary stochastic process with zero mean.*

(1) If $E|X_1|^8 < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-6-\epsilon})$ for some $\epsilon > 0$, then $\kappa_{1,1}(\overline{X}_m, \overline{X}_m^2) = O(m^{-2})$.

(2) If $E|X_1|^{12} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-8-\epsilon})$ for some $\epsilon > 0$, then $\kappa_{2,1}(\overline{X}_m, \overline{X}_m^2) = O(m^{-2})$.

(3) If $E|X_1|^{16} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-10-\epsilon})$ for some $\epsilon > 0$, then $\kappa_{1,2}(\overline{X}_m, \overline{X}_m^2) = O(m^{-3})$ and $\kappa_{3,1}(\overline{X}_m, \overline{X}_m^2) = O(m^{-3})$.

(4) If $E|X_1|^{20} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-12-\epsilon})$ for some $\epsilon > 0$, then $\kappa_{2,2}(\overline{X}_m, \overline{X}_m^2) = O(m^{-3})$ and $\kappa_{4,1}(\overline{X}_m, \overline{X}_m^2) = O(m^{-4})$.

(5) If $E|X_1|^{24} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-14-\epsilon})$ for some $\epsilon > 0$, then $\kappa_{1,3}(\overline{X}_m, \overline{X}_m^2) = O(m^{-4})$, $\kappa_{3,2}(\overline{X}_m, \overline{X}_m^2) = O(m^{-4})$, and $\kappa_{5,1}(\overline{X}_m, \overline{X}_m^2) = O(m^{-5})$.

(6) If $E|X_1|^{28} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-16-\epsilon})$ for some $\epsilon > 0$, then $\kappa_{2,3}(\overline{X}_m, \overline{X}_m^2) = O(m^{-4})$, $\kappa_{4,2}(\overline{X}_m, \overline{X}_m^2) = O(m^{-5})$, and $\kappa_{6,1}(\overline{X}_m, \overline{X}_m^2) = O(m^{-6})$.

(7) If $E|X_1|^{32} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-18-\epsilon})$ for some $\epsilon > 0$, then $\kappa_{3,3}(\overline{X}_m, \overline{X}_m^2) = O(m^{-5})$, $\kappa_{5,2}(\overline{X}_m, \overline{X}_m^2) = O(m^{-6})$, and $\kappa_{7,1}(\overline{X}_m, \overline{X}_m^2) = O(m^{-7})$.

Proof: These results can be obtained from Lemma 2.6. □

Lemma A.8 (Cumulants of $(1/n)\sum_{i=1}^n Y_i^2$). *Assume that $\{X_i : i \geq 1\}$ is a discrete time, strictly stationary stochastic process with zero mean.*

(1) If $E|X_1|^4 < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-5})$, then

$\kappa_1((1/n)\sum_{i=1}^n Y_i^2) = O(b^{-1})$.

(2) If $E|X_1|^{12} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-9})$, then

$\kappa_2((1/n)\sum_{i=1}^n Y_i^2) = O(n^{-1})O(b^{-2})$.

(3) If $E|X_1|^{20} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-13})$, then

$\kappa_3((1/n)\sum_{i=1}^n Y_i^2) = O(n^{-2})O(b^{-3})$.

(4) If $E|X_1|^{28} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-17})$, then

$\kappa_4((1/n)\sum_{i=1}^n Y_i^2) = O(n^{-3})O(b^{-4})$.

Proof: (1) $\kappa_1((1/n)\sum_{i=1}^n Y_i^2) = \kappa_2(Y_1) = O(b^{-1})$.

(2) From Lemma 2.8, the most significant term of $\kappa_2((1/n)\sum_{i=1}^n Y_i^2)$ is

$(1/n)\sum_{i=-\infty}^{\infty} \text{cum}(Y_0^2, Y_i^2)$, which is $(1/n)O(b^{-2})$ form Lemma A.2.

(3) From Lemma 2.8, the most significant term of $\kappa_3((1/n)\sum_{i=1}^n Y_i^2)$ is

$(1/n^2)\sum_{i,j=-\infty}^{\infty} \text{cum}(Y_0^2, Y_i^2, Y_j^2)$, which is $(1/n^2)O(b^{-3})$ from Lemma A.2.

(4) This can be proved in a similar fashion as (2) and (3). □

Lemma A.9 (Mixed cumulants of $\overline{X}_m$ and $(1/n\sum_{i=1}^n Y_i^2)$. *Assume that $\{X_i : i \geq 1\}$ is a discrete time, strictly stationary stochastic process with zero mean.*

(1) If $E|X_1|^8 < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-7})$, then

$\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) = O(n^{-1})O(b^{-2})$.

(2) If $E|X_1|^{12} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-9})$, then

$\kappa_{2,1}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2) = O(n^{-2})O(b^{-2})$.

(3) If $E|X_1|^{16} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-11})$, then

$\kappa_{1,2}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2) = O(n^{-2})O(b^{-3})$, and $\kappa_{3,1}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2) = O(n^{-3})O(b^{-3})$.

(4) If $E|X_1|^{20} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-13})$, then

$\kappa_{2,2}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2) = O(n^{-3})O(b^{-3})$ and $\kappa_{4,1}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2) = O(n^{-4})O(b^{-4})$.

(5) If $E|X_1|^{24} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-15})$, then

$\kappa_{1,3}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2) = O(n^{-3})O(b^{-4})$, $\kappa_{3,2}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2) = O(n^{-4})O(b^{-4})$, and $\kappa_{5,1}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2) = O(n^{-5})O(b^{-5})$.

(6) If $E|X_1|^{28} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-17})$, then

$\kappa_{2,3}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2) = O(n^{-4})O(b^{-4})$, $\kappa_{4,2}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2) = O(n^{-5})O(b^{-5})$, $\kappa_{3,3}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2) = O(n^{-5})O(b^{-5})$, and

$\kappa_{6,1}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2) = O(n^{-6})O(b^{-6})$.

(7) If $E|X_1|^{32} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-19})$, then

$\kappa_{5,2}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2) = O(n^{-6})O(b^{-6})$ and $\kappa_{7,1}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2) = O(n^{-7})O(b^{-7})$.

Proof: (1) From Lemma 2.8, the most significant term of $\kappa_{1,1}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2)$

is $(1/n) \sum_{i=-\infty}^{\infty} \text{cum}(Y_0, Y_i^2)$, which is $(1/n)O(b^{-2})$ form Lemma A.3.

(2) From Lemma 2.8, the most significant term of $\kappa_{2,1}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2)$ is

$(1/n^2) \sum_{i,j=-\infty}^{\infty} \text{cum}(Y_0, Y_i, Y_j^2)$, which is $(1/n^2)O(b^{-2})$ form Lemma A.4.

(3) From Lemma 2.8, the most significant term of $\kappa_{1,2}(\overline{X}_m, (1/n) \sum_{i=1}^{n} Y_i^2)$ is

$(1/n^2) \sum_{i,j=-\infty}^{\infty} \text{cum}(Y_0, Y_i^2, Y_j^2)$, which is $(1/n^2)O(b^{-3})$ form Lemma A.4.

(4)-(7) Can be shown by using the method specified in Remark A.5. □

83

Lemma A.10 (Mixed cumulants of $\overline{X}_m^2$ and $(1/n)\sum_{i=1}^n Y_i^2$). *Assume that $\{X_i : i \geq 1\}$ is a discrete time, strictly stationary stochastic process with zero mean.*

(1) If $E|X_1|^{12} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-9})$, then

$$\kappa_{1,1}(\overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = O(n^{-2})O(b^{-2}).$$

(2) If $E|X_1|^{20} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-13})$, then

$$\kappa_{2,1}(\overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = O(n^{-3})O(b^{-3}) \text{ and } \kappa_{1,2}(\overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) =$$
$$O(n^{-3})O(b^{-3}) + O(n^{-2})O(b^{-4}).$$

Proof: (1) Notice that $\kappa_{1,1}(\overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = \kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)$. The order can be obtained from Lemma A.9(2).

(2) It can be shown that $\kappa_{2,1}(\overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = \kappa_{4,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$4\kappa_3(\overline{X}_m)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 4\kappa_2(\overline{X}_m)\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)$. The desired order follows from Lemma A.9. The second result can be proved by noting

$\kappa_{1,2}(\overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = \kappa_{2,2}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 2\kappa_{1,1}^2(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)$ and by another application of Lemma A.9. $\qquad\square$

Lemma A.11 (Mixed cumulants of $\overline{X}_m$, $\overline{X}_m^2$, and $(1/n)\sum_{i=1}^n Y_i^2$). *Assume that $\{X_i : i \geq 1\}$ is a discrete time, strictly stationary stochastic process with zero mean.*

(1) If $E|X_1|^{16} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-11})$, then

$$\kappa_{1,1,1}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = O(n^{-2})O(b^{-3}).$$

(2) If $E|X_1|^{20} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-13})$, then

$$\kappa_{2,1,1}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = O(n^{-3})O(b^{-3}).$$

(3) If $E|X_1|^{24} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-15})$, then

$\kappa_{1,2,1}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = O(n^{-3})O(b^{-4})$, $\kappa_{1,1,2}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = O(n^{-3})O(b^{-4})$, and $\kappa_{3,1,1}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = O(n^{-4})O(b^{-4})$.

(4) If $E|X_1|^{28} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-17})$, then

$\kappa_{2,2,1}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = O(n^{-4})O(b^{-4})$, $\kappa_{2,1,2}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = O(n^{-4})O(b^{-4})$, and $\kappa_{4,1,1}(\overline{X}_n, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = O(n^{-5})O(b^{-5})$.

(5) If $E|X_1|^{32} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-19})$, then

$\kappa_{3,2,1}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = O(n^{-5})O(b^{-5})$ and $\kappa_{3,1,.}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2)$
$= O(n^{-5})O(b^{-5})$.

Proof: For each of the desired results, we first derive an identity and then derive the order from Lemma A.9. Those identities are as follows.

(1) $\kappa_{1,1,1}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = \kappa_{3,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$
$2\kappa_2(\overline{X}_m)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)$.

(2) $\kappa_{2,1,1}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = \kappa_{4,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$
$2\kappa_3(\overline{X}_m)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 4\kappa_2(\overline{X}_m)\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)$.

(3) $\kappa_{1,2,1}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = \kappa_{5,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$
$4\kappa_4(\overline{X}_m)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 8\kappa_2(\overline{X}_m)\kappa_{3,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$
$8\kappa_3(\overline{X}_m)\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 8\kappa_2^2(\overline{X}_m)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)$;

$\kappa_{1,1,2}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = \kappa_{3,2}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)$
$+ 4\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$
$2\kappa_2(\overline{X}_m)\kappa_{1,2}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)$; and

$\kappa_{3,1,1}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = \kappa_{5,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$
$2\kappa_4(\overline{X}_m)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 6\kappa_2(\overline{X}_m)\kappa_{3,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$
$6\kappa_3(\overline{X}_m)\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)$.

(4) $\kappa_{2,2,1}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = \kappa_{6,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)$

$+ 12\kappa_2(\overline{X}_m)\kappa_{4,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 4\kappa_5(\overline{X}_m)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$16\kappa_3(\overline{X}_m)\kappa_{3,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 12\kappa_4(\overline{X}_m)\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$24\kappa_2^2(\overline{X}_m)\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 24\kappa_3(\overline{X}_m)\kappa_2(\overline{X}_m)\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2);$

$\kappa_{2,1,2}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = \kappa_{4,2}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$4\kappa_{3,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$2\kappa_3(\overline{X}_m)\kappa_{1,2}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 4\kappa_2(\overline{X}_m)\kappa_{2,2}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$4\kappa_{2,1}^2(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2);$ and

$\kappa_{4,1,1}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = \kappa_{6,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$8\kappa_2(\overline{X}_m)\kappa_{4,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 2\kappa_5(\overline{X}_m)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$12\kappa_3(\overline{X}_m)\kappa_{3,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 8\kappa_4(\overline{X}_m)\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2).$

(5) $\kappa_{3,2,1}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = \kappa_{7,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$16\kappa_2(\overline{X}_m)\kappa_{5,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 4\kappa_6(\overline{X}_m)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$28\kappa_3(\overline{X}_m)\kappa_{4,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 16\kappa_5(\overline{X}_m)\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$28\kappa_4(\overline{X}_m)\kappa_{3,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 48\kappa_2^2(\overline{X}_m)\kappa_{3,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$24\kappa_3^2(\overline{X}_m)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 32\kappa_4(\overline{X}_m)\kappa_2(\overline{X}_m)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$96\kappa_3(\overline{X}_m)\kappa_2(\overline{X}_m)\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2);$ and

$\kappa_{3,1,2}(\overline{X}_m, \overline{X}_m^2, (1/n)\sum_{i=1}^n Y_i^2) = \kappa_{5,2}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$6\kappa_2(\overline{X}_m)\kappa_{3,2}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) + 6\kappa_3(\overline{X}_m)\kappa_{2,2}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$4\kappa_{4,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)\kappa_{1,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$12\kappa_{3,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2)\kappa_{2,1}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$2\kappa_4(\overline{X}_m)\kappa_{1,2}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2) +$

$3\kappa_2^2(\overline{X}_m)\kappa_{1,2}(\overline{X}_m, (1/n)\sum_{i=1}^n Y_i^2).$ $\qquad\square$

Lemma A.12 (Cumulants of $\Delta_{n,b}$). *Assume that $\{X_i : i \geq 1\}$ is a discrete time, strictly stationary stochastic process with zero mean.*

(1) If $E|X_1|^4 < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-5})$, then $\kappa_1(\Delta_{n,b}) = O(b^{-1}) + O(n^{-1})$

(2) If $E|X_1|^{12} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-9})$, then $\kappa_2(\Delta_{n,b}) = O(n^{-1})$.

(3) If $E|X_1|^{20} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-13})$, then $\kappa_3(\Delta_{n,b}) = O(n^{-2})$

Proof: Recall that $\Delta_{n,b} \equiv V_{n,b}/(\sigma_m^2/b) - 1$ and $V_{n,b} = (1/n)\sum_{i=1}^{n} Y_i^2 - \overline{X}_m^2$, then

(1) $\kappa_1(\Delta_{n,b}) = E[V_{n,b}/(\sigma_m^2/b) - 1] = (b/\sigma_m^2)E((1/n)\sum_{i=1}^{n} Y_i^2 - 1) - (b/\sigma_m^2)E\overline{X}_m^2$

$= (b/\sigma_m^2)E(Y_1^2 - 1) - (b/\sigma_m^2)E\overline{X}_m^2 = O(b^{-1}) + O(n^{-1})$.

(2) $\kappa_2(\Delta_{n,b}) = (b/\sigma_m^2)^2\kappa_2[(1/n)\sum_{i=1}^{n} Y_i^2 - \overline{X}_m^2] = (b/\sigma_m^2)^2[O(n^{-1})O(b^{-2}) + O(n^{-2})O(b^{-2}) + O(n^{-2})O(b^{-2})] = O(n^{-1})$, where the orders are obtained from Corollary A.6, Lemma A.8, and Lemma A.10.

(3) $\kappa_3(\Delta_{n,b}) = (b/\sigma_m^2)^3\kappa_3[(1/n)\sum_{i=1}^{n} Y_i^2 - \overline{X}_m^2] = (b/\sigma_m^2)^3[O(n^{-2})O(b^{-3})] = O(n^{-2})$, where the orders are obtained from Corollary A.6, Lemma A.8, and Lemma A.10. $\square$

Lemma A.13 (Mixed cumulants of $\overline{X}_m$ and $\Delta_{n,b}$). *Assume that $\{X_i : i \geq 1\}$ is a discrete time, strictly stationary stochastic process with zero mean.*

(1) If $E|X_1|^8 < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-7})$, then $\kappa_{1,1}(\Delta_{n,b}, \overline{X}_m) = O(n^{-1})O(b^{-1})$.

(2) If $E|X_1|^{16} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-11})$, then $\kappa_{2,1}(\Delta_{n,b}, \overline{X}_m) = O(n^{-2})O(b^{-1})$.

87

(3) If $E|X_1|^{12} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-9})$, then

$$\kappa_{1,2}(\Delta_{n,b}, \overline{X}_m) = O(n^{-2})O(b^{-1}).$$

(4) If $E|X_1|^{24} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-15})$, then

$$\kappa_{3,1}(\Delta_{n,b}, \overline{X}_m) = O(n^{-3})O(b^{-1}).$$

(5) If $E|X_1|^{20} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-13})$, then

$$\kappa_{2,2}(\Delta_{n,b}, \overline{X}_m) = O(n^{-3})O(b^{-1}).$$

(6) If $E|X_1|^{16} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-11})$, then

$$\kappa_{1,3}(\Delta_{n,b}, \overline{X}_m) = O(n^{-3})O(b^{-2}).$$

(7) If $E|X_1|^{28} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-17})$, then

$$\kappa_{3,2}(\Delta_{n,b}, \overline{X}_m) = O(n^{-4})O(b^{-1}).$$

(8) If $E|X_1|^{24} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-15})$, then

$$\kappa_{2,3}(\Delta_{n,b}, \overline{X}_m) = O(n^{-4})O(b^{-2}).$$

(9) If $E|X_1|^{20} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-13})$, then

$$\kappa_{1,4}(\Delta_{n,b}, \overline{X}_m) = O(n^{-4})O(b^{-3}).$$

(10) If $E|X_1|^{32} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-19})$, then

$$\kappa_{3,3}(\Delta_{n,b}, \overline{X}_m) = O(n^{-5})O(b^{-2}).$$

(11) If $E|X_1|^{28} < \infty$, and the sequence is mixing with $\alpha_n = O(n^{-17})$, then

$$\kappa_{2,4}(\Delta_{n,b}, \overline{X}_m) = O(n^{-5})O(b^{-3}).$$

Proof: Recall that $\Delta_{n,b} = V_{n,b}/(\sigma_m^2/b) - 1$ and $V_{n,b} = (1/n)\sum_{i=1}^{n} Y_i^2 - \overline{X}_m^2$, then

(1) $\kappa_{1,1}(\Delta_{n,b}, \overline{X}_m) = (b/\sigma_m^2)[\kappa_{1,1}((1/n)\sum_{i=1}^{n} Y_i^2, \overline{X}_m) + \kappa_{1,1}(\overline{X}_m^2, \overline{X}_m)] =$

$O(n^{-1})O(b^{-1})$, where the orders are obtained from Corollary A.7 and Lemma A.9.

(2) $\kappa_{2,1}(\Delta_{n,b}, \overline{X}_m) = (b/\sigma_m^2)^2\kappa_{1,1,1}[(1/n)\sum_{i=1}^{n} Y_i^2 - \overline{X}_m^2, (1/n)\sum_{i=1}^{n} Y_i^2 - \overline{X}_m^2, \overline{X}_m]$

$= O(n^{-2})O(b^{-1})$, where the orders are obtained from Corollary A.7, Lemma A.9.

and Lemma A.11.

(3) $\kappa_{1,2}(\Delta_{n,b}, \overline{X}_m) = (b/\sigma_m^2)\kappa_{1,1,1}[(1/n)\sum_{i=1}^n Y_i^2 - \overline{X}_m^2, \overline{X}_m, \overline{X}_m] = O(n^{-2})O(b^{-1})$,

where the orders are obtained from Corollary A.7 and Lemma A.9.

(4) $\kappa_{3,1}(\Delta_{n,b}, \overline{X}_m) = (b/\sigma_m^2)^3\kappa_{3,1}[(1/n)\sum_{i=1}^n Y_i^2 - \overline{X}_m^2, \overline{X}_m]$

$= (b/\sigma_m^2)^3 O(n^{-3})O(b^{-4}) = O(n^{-3}b^{-1})$, where the orders are obtained from Corollary A.7, Lemma A.9, and Lemma A.11.

(5) $\kappa_{2,2}(\Delta_{n,b}, \overline{X}_m) = (b/\sigma_m^2)^2\kappa_{2,2}[(1/n)\sum_{i=1}^n Y_i^2 - \overline{X}_m^2, \overline{X}_m]$

$= (b/\sigma_m^2)^2 O(n^{-3})O(b^{-3}) = O(n^{-3})O(b^{-1})$, where the orders are obtained from Corollary A.7, Lemma A.9, and Lemma A.11.

(6) $\kappa_{1,3}(\Delta_{n,b}, \overline{X}_m) = (b/\sigma_m^2)\kappa_{1,3}[(1/n)\sum_{i=1}^n Y_i^2 - \overline{X}_m^2, \overline{X}_m]$

$= (b/\sigma_m^2)^2 O(n^{-3})O(b^{-3}) = O(n^{-3})O(b^{-2})$, where the orders are obtained from Corollary A.7, Lemma A.9, and Lemma A.11.

(7) $\kappa_{3,2}(\Delta_{n,b}, \overline{X}_m) = (b/\sigma_m^2)^3\kappa_{3,2}[(1/n)\sum_{i=1}^n Y_i^2 - \overline{X}_m^2, \overline{X}_m]$

$= (b/\sigma_m^2)^3 O(n^{-4})O(b^{-4}) = O(n^{-4})O(b^{-1})$, where the orders are obtained from Corollary A.7, Lemma A.9, and Lemma A.11.

(8) $\kappa_{2,3}(\Delta_{n,b}, \overline{X}_m) = (b/\sigma_m^2)^2\kappa_{2,3}[(1/n)\sum_{i=1}^n Y_i^2 - \overline{X}_m^2, \overline{X}_m]$

$= (b/\sigma_m^2)^2 O(n^{-4})O(b^{-4}) = O(n^{-4})O(b^{-2})$, where the orders are obtained from Corollary A.7, Lemma A.9, and Lemma A.11.

(9) $\kappa_{1,4}(\Delta_{n,b}, \overline{X}_m) = (b/\sigma_m^2)\kappa_{1,4}[(1/n)\sum_{i=1}^n Y_i^2 - \overline{X}_m^2, \overline{X}_m]$

$= (b/\sigma_m^2)^2 O(n^{-4})O(b^{-4}) = O(n^{-4})O(b^{-3})$, where the orders are obtained from Corollary A.7, Lemma A.9, and Lemma A.11.

(10) $\kappa_{3,3}(\Delta_{n,b}, \overline{X}_m) = (b/\sigma_m^2)^3\kappa_{3,3}[(1/n)\sum_{i=1}^n Y_i^2 - \overline{X}_m^2, \overline{X}_m]$

$= (b/\sigma_m^2)^3 O(n^{-5})O(b^{-5}) = O(n^{-5})O(b^{-2})$, where the orders are obtained from Corollary A.7, Lemma A.9, and Lemma A.11.

(11) $\kappa_{2,4}(\Delta_{n,b}, \overline{X}_m) = (b/\sigma_m^2)^2\kappa_{2,4}[(1/n)\sum_{i=1}^n Y_i^2 - \overline{X}_m^2, \overline{X}_m]$

$= (b/\sigma_m^2)^2 O(n^{-5}) O(b^{-5}) = O(n^{-5}) O(b^{-3})$, where the orders are obtained from Corollary A.7, Lemma A.9, and Lemma A.11. $\qquad\square$

Proof of Proposition 3.4: Recall that $\{X_i\}$ is zero mean. From equation (3.10),

$t_{n,b} = (m^{1/2}/\sigma_m)\overline{X}_m$

$\{1 - (1/2)\Delta_{n,b} + (3/8)\Delta_{n,b}^2 - (5/16)\Delta_{n,b}^3 + \cdots\}.$

(1) $\kappa_1(t_{n,b}) = (m^{1/2}/\sigma_m)E[\overline{X}_m\{1 - (1/2)\Delta_{n,b} + (3/8)\Delta_{n,b}^2 - (5/16)\Delta_{n,b}^3 + \cdots\}] = (m^{1/2}/\sigma_m)E(\overline{X}_m) - (1/2)(m^{1/2}/\sigma_m)E[\overline{X}_m\Delta_{n,b}] + \cdots$. The first term on the right hand side is 0, the second term is $O(n^{-1/2}b^{-1/2})$ and has the most significant term $-(1/2)(\sqrt{m}b/\sigma_m^3)\kappa_{1,1}(\overline{X}_m,(1/n)\sum_{i=1}^n Y_i^2)$, and the remainder terms are of $o(n^{-1/2}b^{-1/2})$.

(2) $\kappa_2(t_{n,b}) = (m/\sigma_m^2)\kappa_2(\overline{X}_m) - (m/\sigma_m^2)\kappa_{1,1}(\overline{X}_m,\Delta_{n,b}) + (1/4)(m/\sigma_m^2)\kappa_2(\Delta_{n,b}) + \ldots$ The desired result then follows from Lemmas A.12 and A.13.

(3) Note that $\kappa_3(t_{n,b}) = (m^{3/2}/\sigma_m^3)\kappa_3(\overline{X}_m) - (3/2)(m^{3/2}/\sigma_m^3)\kappa_{2,1}(\overline{X}_m,\Delta_{n,b}) + (3/4)(m^{3/2}/\sigma_m^3)\kappa_{1,2}(\overline{X}_m,\Delta_{n,b}) - (1/8)(m^{3/2}/\sigma_m^3)\kappa_3(\Delta_{n,b}) + \ldots$. The desired result tehn follows from Lemmas A.12 and A.13.

(4) Note that $\kappa_3(t_{n,b}) = (m^2/\sigma_r^4)\kappa_4(\overline{X}_m) - 2(m^2/\sigma_m^4)\kappa_{3,1}(\overline{X}_m,\Delta_{n,b}) + (3/2)(m^2/\sigma_m^4)\kappa_{2,2}(\overline{X}_m,\Delta_{n,b}) - (1/2)(m^2/\sigma_m^4)\kappa_{1,3}(\overline{X}_m,\Delta_{n,b}) + (1/16)(m^2/\sigma_m^4)\kappa_4(\Delta_{n,b}) + \ldots$. The desired result then follows from Lemmas A.12 and A.13. $\qquad\square$

References

[1] Anderson, T. W. (1971). *The Statistical Analysis of Time Series*. John Wiley & Sons, New York.

[2] Bhattacharya, R. N., and Ghosh, J. K. (1978). On the validity of formal Edgeworth expansion. *Ann. Statist.* **6** 434–451.

[3] Bhattacharya, R. N., and Ranga Rao, R. (1976). *Formal Approximation and Asymptotic Expansions*. Wiley, New York.

[4] Billingsley, P. (1968). *Convergence of Probability Measures*. John Wiley & Sons, New York.

[5] Billingsley, P. (1979). *Probability and Measure*. John Wiley & Sons, New York.

[6] Brillinger, D. R. (1973). Estimation of the mean of a stationary time series by sampling. *J. Appl. Prob.* **10** 419–431.

[7] Brillinger, D. R. (1981). *Time Series Analysis: Data Analysis and Theory*. Holt, Rinehart & Winston, New York.

[8] Chung, K. L. (1974). *A Course in Probability Theory*, 2nd ed. Academic Press, New York.

[9] Cornish, E. A., and Fisher, R. A. (1937). Moments and cumulants in the specification of distributions. *Rev. Int. Statist. Inst.* **5** 307–320.

[10] Cramér, H. (1946). *Mathematical Methods of Statistics.* Princeton University Press, Princeton, New Jersey.

[11] Efron, B. (1978). Nonparametric standard errors and confidence intervals. *Canad. J. Statist.* **9** 139-172.

[12] Glynn, P. W. (1982). Asymptotic theory for nonparametric confidence intervals. Technical Report, No. 19, Dept. of Operations Research, Stanford University.

[13] Hall, P. (1983). Inverting an Edgeworth expansion. *Ann. Statist.* **39** 1264-1273.

[14] Iglehart, D. L., unpublished class notes.

[15] Johnson, N. J. (1978). Modified *t*-tests and confidence intervals for asymmetrical populations. *J. Amer. Statist. Assoc.* **73** 536-544.

[16] Kendall, M. G., and Stuart, A. (1977). *The Advanced Theory of Statistics,* 4th ed., vol. 1. Macmillan, New York.

[17] Lamperti, J. (1977). *Stochastic Processes: A Survey of the Mathematical Theory.* Springer-Verlag, New York.

[18] **Law, A. (1983). Statistical analysis of simulation output data.** *Oper. Res.* **31** 983-1029.

[19] Law, A., and Kelton, D. (1984). Confidence intervals for steady state simulations: I. A Survey of fixed sample size procedures. *Oper. Res.* **32** 1221-1239.

[20] Lukacs, E. (1970). *Characteristic Functions,* 2nd ed. Hafner, New York.

[21] Prakasa Rao, B. L. S. (1987). *Asymptotic Theory of Statistical Inference*. John Wiley & Sons, New York.

[22] Rosenblatt, M. (1985). *Stationary Sequences and Random Fields*, Birkhäuser, Boston.

[23] Schmeiser, B. (1982). Batch size effects in the analysis of simulation output. *Op. Resch.* **30** 556–568.

[24] Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons. New York.

[25] Titus, B. D. (1985). Modified confidence intervals for the mean of an autoregressive process. Technical Report, No. 34, Dept. of Operations Research. Stanford University.

[26] Wallace, D. L. (1958) Asymptotic approximations of distribution functions. *Ann. Math. Statist.* **29** 635–654.

[27] Withers, C. S. (1983). Expansions for the distribution and quantiles of a regular functional of the empirical distribution. *Ann. Statist.* **11** 577–587.

MASTER COPY   -   FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS			
Unclassified				

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
Technical Report No. 37	ARO 25839.12.MA

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Dept. of Operations Research		U. S. Army Research Office

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
Stanford, CA 94305-4022	P. O. Box 12211 Research Triangle Park, NC 27709-2211

8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
U. S. Army Research Office		DAAL03-88-K-0063

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
P. O. Box 12211 Research Triangle Park, NC 27709-2211	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO

11. TITLE (Include Security Classification)

Small Sample Theory for Steady State Confidence Intervals

12. PERSONAL AUTHOR(S) Chia-Hon Chien

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM ____ TO ____	June 1989	93

16. SUPPLEMENTARY NOTATION The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

17.	COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	simulation, steady-state, nonparametric confidence intervals,
			Edgeworth expansions, stationary processes

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Please see other side.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted
All other editions are obsolete

The goal of this dissertation is to develop a nonparametric method for obtaining a confidence interval for the mean of a stationary sequence. As indicated in the literature, nonparametric confidence intervals in practice often have undesirable small-sample asymmetry and coverage characteristics. These phenomena are partially due to the fact that the third and fourth cumulants of the point estimator for the stationary mean, unlike those of the standard normal random variable, are not zero. We will apply Edgeworth and Cornish-Fisher expansions to obtain asymptotic expansions for the errors associated with confidence intervals. The analysis isolates various elements that contribute to errors and makes it possible for us to estimate each element and hopefully correct the errors to a smaller order. We will use Glynn's method to develop first and second order pivots for the confidence intervals. Furthermore, these procedures also improve the asymptotic order of confidence interval accuracy.